Thermal Adaptation of Enzymes: Impacts of Conformational Shifts on Catalytic Activation Energy and Optimum Temperature

Irene Maffucci, Damien Laage, Fabio Sterpone, Guillaume Stirnemann

Thermal adaptation of enzymes: impacts of conformational shifts on catalytic activation energy and optimum temperature

Irene Maffucci[1,2,‡], Damien Laage[1*], Fabio Sterpone[2*], and Guillaume Stirnemann[2*]

[1]*PASTEUR, Département de chimie, École Normale Supérieure, PSL University, Sorbonne Université, CNRS, 24 rue Lhomond, 75005 Paris, France*

[2]*CNRS Laboratoire de Biochimie Théorique, Institut de Biologie Physico-Chimique, PSL University, Université de Paris, 13 rue Pierre et Marie Curie, 75005, Paris, France*

[*]*To whom corresponce should be addressed: damien.laage@ens.psl.eu, sterpone@ibpc.fr, stirnemann@ibpc.fr*

[‡] *Present address: Université de Technologie de Compiègne, UPJV, CNRS, Enzyme and Cell Engineering, Centre de recherche Royallieu - CS 60319 - 60203 Cedex, France*

Thermal adaptation of enzymes is essential for both living organism development in extreme conditions and efficient biocatalytic applications. However, the molecular mechanisms leading to a shift in catalytic activity optimum temperatures remain unclear, and there is increasing experimental evidence that thermal adaptation involves complex changes in both structural and reactive properties. Here we apply a combination of enhanced protein conformational sampling with an explicit chemical reaction description to mesophilic and thermophilic homologs of the dihydrofolate reductase enzyme, and obtain a quantitative description of the stability and catalytic activity shifts between homologs. In contrast with pictures focusing on protein flexibility and dynamics, we reveal the key role played by temperature-induced shifts in protein conformational distributions; we show that while the homologs' reaction free energies are similar, the striking discrepancy between their activation energies is caused by their different conformational changes with temperature. We propose an analytic model

combining catalytic activity and structural stability which quantitatively predicts the shift in homologs' optimum temperatures, and we show that this general model provides a molecular explanation of changes in optimum temperatures for several other enzymes.

Introduction

Life on Earth spans a $\approx 140°$C temperature-range.[1] In order to guarantee growth and reproduction of organisms thriving in harsh conditions,[2,3] evolution has thus led, through mutations, to conserved enzyme families that catalyze similar chemical reactions but in radically different thermal environments, from the very cold regime for psychrophiles, to the ambient regime in mesophiles, and up to very high temperatures in thermophilic organisms.[4]

The first obvious effect of temperature is to affect protein stability. Protein folded structures are usually very sensitive to environmental conditions: small changes in temperature but also in pH or the presence of osmolytes can lead to protein denaturation, and thus a loss of function. The molecular mechanisms by which mutations have led to temperature adaptation by shifting the protein melting temperature T_m (defined as the temperature at which half the proteins are unfolded) have been intensively studied.[5] For example, the matrix of thermophilic proteins is often found to be more rigid than that of mesophiles.

However, in addition to structural stability, another major requirement for enzymes is to retain their catalytic activity. While a reduced structural flexibility can increase thermal stability,[5,6] it also hinders the conformational rearrangements necessary for catalysis, leading to a significant activity loss and an activity-stability trade-off. Experimentally, enzymatic activity is measured to increase with temperature until an optimum activity temperature T_{opt} above which it drops abruptly. Well below this optimum temperature, the temperature dependence of the catalytic rate is Arrhenian, i.e., the enzymatic rate constant follows the Arrhenius law with a constant activation energy.

Several models have been proposed to explain the presence and the location of the activity optimum,[5,7,8] involving reactivity and structural stability features. A straightforward explanation is that it could arise from protein denaturation that would lead to a sudden drop in the Arrhenian dependence of the rate constant increase with temperature; in that case, T_{opt} would thus be expected to closely follow the protein melting temperature T_m. But strikingly, experiments have shown that the $T_m - T_{opt}$ shift is not constant among homologs.[9,10] Another recent proposal is that an activity optimum could arise from reaction activation heat capacity terms, causing a deviation from the simple Arrhenius law with a constant activation energy;[11,12] however, the typically small difference between the T_m and T_{opt} temperatures suggests that unfolding should play a role in determining the activity optimum, and the effect of melting thus cannot be neglected.[13]

Here we address two major aspects of thermal adaptation that have remained elusive. The first one focuses on the molecular factors which determine the optimum activity temperature, and how much it differs from the melting temperature. A drop in activity well below melting has been interpreted in terms of an equilibrium between catalytically active and inactive forms of the enzyme,[14–16] but a direct quantification of conformational ensembles with distinct activity, which is not easily accessed either in the experiments or in simulations, is missing. The second one considers temperatures well below T_m and deals with the Arrhenius activation energy, which determines how sensitive the reaction rate constant is vis-a-vis a temperature change. Enzymes adapted to lower temperatures typically exhibit smaller reaction activation energies, which has been suggested to avoid a fast activity drop with decreasing temperatures.[17] However, this difference among homologs sharing identical active sites is surprising,[18] especially since the free energy barriers are often similar. A recent proposal suggested[19,20] that different surface group flexibilities among homologs could tune the entropic and enthalpic contributions, and explain the difference in their activation energies. However, the molecular origin for changes in the surface group properties between the enzyme reactant and transition state structures remains to be elucidated.

Enzymes are known to typically sample a broad distribution of conformations, possibly having different catalytic activities.[21] This is expected to be critical as the temperature increases towards the melting point and a key issue is to determine how the equilibrium between multiple conformations differs among homologs. Molecular simulations have been shown to be an ideally suited tool to obtain the required molecular insight and have already proved successful in the calculation of melting temperatures[22,23] and rate constants of enzyme-catalyzed reactions.[24–26] An often overlooked aspect in computational studies of enzyme catalysis is that while characterizing this conformational distribution remains very difficult experimentally, recent developments in enhanced sampling techniques now give access to a proper sampling of these conformations, even though it remains challenging and requires an important simulation effort. These may not be captured in regular brute force simulations where the protein can be trapped in conformations irrelevant at high temperatures, even when sampling over many short trajectories, as routinely done.

Here, we address these critical aspects by combining all-atom molecular dynamics simulations with enhanced conformational sampling and a powerful coupled valence-bond state description of chemical reactions. The model systems that we have selected are the (monomeric) mesophilic *Escherichia coli* (Ec) and the (dimeric) thermophilic *Thermotoga maritima* (Tm) homologs of the paradigm dihydrofolate reductase (DHFR) enzyme[27] (Fig. 1), for which a large body of experimental and simulation data are available, and which exhibit dramatic differences in the catalytic rate temperature dependence as mentioned above. The chemical reaction catalyzed by DHFR is a hydride transfer which reduces 7,8-dihydrofolate (FOL) into 5,6,7,8-tetrahydrofolate (THF) with NADPH as cofactor (Fig. 1c). While this reaction is not rate-limiting for the overall catalytic process, its rate is experimentally accessible via pre-steady state kinetics. Experimental measurements[28–31] have shown that at temperatures well below the protein melting point, both Ec and TmDHFR exhibit Arrhenian behaviors, with three major differences: first, catalytic activity is

much larger in Ec than in TmDHFR; second, the activation energy, i.e. the rate constant temperature dependence, is significantly smaller in Ec than in TmDHFR; and third, the temperature of maximum activity for TmDHFR is located very close to (3 K below) its melting point T_m, while it is 16 K below T_m for EcDHFR. The reduced activity in TmDHFR has been intensively studied and has been attributed to a key structural change[32–37] in a loop that is crucial for catalysis.[26,27,38–40] While these conclusions are indeed confirmed by our present results, we will rather focus on the other, so far unexplained aspects pertaining to the large differences in the catalytic rate temperature dependences and to the shift between their optimum and melting temperatures. We show that our simulations can quantitatively reproduce all the differences in the temperature dependence of the catalytic rates between the two selected DHFR homologs. They further reveal the molecular origin for the much reduced activation energy in a mesophilic enzyme as compared to its thermophilic counterpart, and show that it is caused by a key conformational change preceding thermal denaturation. We build an analytic model that predicts the temperature dependence of the catalytic rate based on both protein stability and the reaction activation energy, and explains the large difference between optimal and melting temperatures in the mesophile. We finally show that this general model provides a molecular explanation of changes in optimum temperatures for a broad range of enzymes.

Conformational sampling on a wide temperature range

Because free energy barriers of the catalyzed reaction can sensitively depend on the enzyme conformation[21], our computational scheme combines extensive conformational sampling with a reactive force field to calculate reaction free energy barriers. Following prior extensive studies,[41–45] we calculate the hydride transfer reaction free energy barrier $\Delta G^{\ddagger}$ with an empirical valence bond (EVB)[41,45] approach (see Methods and Supplementary Information). The electronic structure of the substrate and cofactor is modeled as a linear combination of two coupled resonance states, each

described by a classical force-field derived from quantum calculations. A major advantage of this method over traditional QM/MM approaches lies in its computational efficiency that allows for an extensive sampling of protein/substrate configurations (a total of $>$ 160 ns of simulations at the reaction transition state was required here), which we will show to be critical.

An important aspect for both EVB and traditional QM/MM approaches is that, while a lot of effort is invested in computing the reaction free energy barrier, it is equally crucial, and often neglected, to sample the enzyme reactant conformations as accurately as possible. Indeed, several conformations of the enzyme, even structurally close to the native state, may be present at a given temperature and exhibit different reactivities. Experimentally, the measured catalytic rate would be averaged over this ensemble of structures. Repeating EVB calculations several times but starting from the same initial structure would therefore not capture such aspect. This is of particular importance as temperature increases, which can critically alter the protein structure. Several experimental studies have clearly identified the presence of a significant population of non-native protein conformations for EcDHFR at temperatures well below melting.[46–48] In particular, these intermediates involve large changes in the Met20 loop[49], which plays an important role during catalysis.[26,27,38–40] These results were confirmed by our own very recent study[37] which evidenced a significant opening of this loop in EcDHFR below melting, whereas it is predominantly closed around the reactants in the native state at ambient temperature.

Unfortunately, sampling multiple protein conformations is not routinely accessible by conventional, brute force simulations, because these conformations are usually separated by large free energy barriers, leading to interconversion timescales that largely exceed the current accessible timescale of the simulations. For EcDHFR, NMR experiments have shown that the exchange between Met20 conformations typically occurs on timescales larger than a millisecond.[50] Within the framework of Kramers theory, MD simulations in implicit solvent have shown that this timescale

is due to a combination of a noticeable free energy barrier with a slow diffusion coefficient along the open/closed coordinate.[51]

Our brute force test MD simulations performed at the experimental melting temperatures evidence these challenges. For each system, we started from the crystal structure (monomeric state for EcDHFR and the dimer for TmDHFR, see Methods and Supplementary Information) that was solvated and propagated for 1 μs at T_m^{exp}, respectively 326 K[48] and 356 K.[30] As expected, no noticeable change of the proteins structure was observed on this timescale with this simulation setup, with atomistic fluctuations of the protein backbone very close to those observed at 300 K (see Supplementary Information Section II). In contrast, experimental data show that at this temperature, half of the enzymes should populate non-native conformations, which is clearly not the case in the brute force simulations. Our own enhanced sampling simulations[37] suggest that the enzyme structures around melting are significantly different from the native state, and we will now show that a successful strategy is to utilize the conformations extracted from enhanced sampling trajectories.

A straightforward approach to thermally unfold proteins in the simulations is to propagate trajectories at extremely high and unphysical temperatures,[52,53] but with no direct and easy correspondance established between the simulation temperature and its equivalent in the experiments.[54] In order to facilitate the sampling of non-native but experimentally-relevant conformations, without an a priori knowledge or any assumption made on the corresponding reaction coordinates, which can be very complex, we ran solute-tempering Hamiltonian replica exchange (REST2[55,56]) simulations with an all-atom description of the protein and its solvent. Each of the 24 replicas was propagated for 0.5 μs at the same physical temperature, but with a different rescaling of the protein potential energy term; a mean-field approximation with a corresponding state formalism is then used to recover an effective, corresponding temperature, called the temperature T in the following.

A key advantage of this approach is that protein stability curves can be recovered, as shown

in our own recent work.[37] Another important aspect is that conformations at different temperatures can be generated, from the ambient regime to well above melting. A systematic difference between experimental and computed absolute T_m values was observed,[37] as already reported for other systems.[22,57] This is likely due to a combination of force-field limitations, finite sampling, and to the temperature reconstruction method. In particular, protein force-fields are usually calibrated on ambient-temperature data or calculations, and are not optimized for significantly different temperatures. Other approaches using temperature replica exchange on much smaller polypeptides, which are not applicable to our large systems but which do not use our indirect temperature reconstruction scheme, have shown that the simulation melting temperature is highly protein and solvent forcefield dependent, and can differ from the experimental T_m value by more than 100 K.[52,58–61]

However, quite strikingly, our simulations very well reproduced[37] the experimental melting temperature shift between Ec and TmDHFR ($\Delta T_m^{sim} = 28$ K vs $\Delta T_m^{exp} = 30$ K,[30,48]) and were therefore able to account for the increased thermal stability of TmDHFR. They also showed significant conformational changes in the Met20 loop of EcDHFR,[37] in line with the experimental results. As we show now, these loop conformational changes have a dramatic impact on the catalytic rate temperature dependence.

Catalytic activity

We now describe the results of our reaction free energy barrier calculations and identify the enzyme structural features which impact the barrier. This will be essential to determine the average rate constant and how it changes with temperature.

For a series of effective temperatures from our REST2 simulations, we ran EVB simulations at the same temperatures to obtain reaction free energy barrier $\Delta G^{\ddagger}$ values in Fig. 2a for 10 configurations at each temperature (see Methods and Supplementary Information). For the TmDHFR

dimer, we focused on the reaction in one of the domains only. We find that on average, $\Delta G^{\ddagger}$ is larger in TmDHFR than in EcDHFR over the investigated temperature range, in agreement with the experimentally measured activity difference between Tm and EcDHFR.[62] Moreover, what was not accessible experimentally and is revealed by our simulations is the increasing spread in $\Delta G^{\ddagger}$ with increasing temperature, suggesting an increasing conformational heterogeneity. It also shows that Ec and TmDHFR have strikingly different changes in the average $\langle \Delta G^{\ddagger} \rangle$ with temperature: while $\langle \Delta G^{\ddagger} \rangle$ is approximately constant in TmDHFR (in agreement with recent calculations[36] on a narrower temperature range), it exhibits a sharp increase around 330 K in EcDHFR, i.e. $\approx$70 K below the simulated T_m.

We now show that this $\langle \Delta G^{\ddagger} \rangle$ temperature dependence is correlated with the Met20 loop conformation in the folded state (Fig. 2b). We stress that these conformational changes take place without global unfolding and occur much below the melting temperature. Several prior studies[26,27,38–40] had stressed the important role played by this loop during the chemical step, and our present results now provide a quantitative connection between loop conformation and chemical barrier height. (We pause to note that the impact of loop conformations on catalytic activity is not specific to DHFR, as shown by a recent study of another enzymatic reaction.[63]) Following our previous study focusing on the effect of temperature on the enzymes structure,[37] we monitor the Met20 loop opening and closing via the Met20 – αC helix distance (see Fig. 1 and Supplementary Section III.1). Figure 2b shows that, at ambient temperatures, this distance is short in EcDHFR, i.e. the loop is closed, while it is open in TmDHFR, in agreement with crystallographic structures (EcDHFR: 1RX2 and TmDHFR: 1D1G) and prior studies.[27,34,36] When temperature increases, the loop exhibits a sudden change in EcDHFR from closed to open conformations at 330 K, while in TmDHFR it remains in the open configuration at all temperatures.

The Met20 loop has been suggested to play two important roles during the catalyzed hy-

dride transfer:[27,40,51,64] i) keeping the folate substrate next to the cofactor and ii) protecting the substrate reactive N5 site (Fig. 1c) from the polar aqueous solvent.[65] These two aspects can be best understood using a theoretical model[66] that established that hydride-(and proton-)transfer free energy barriers sensitively depend on the rearrangement of two coordinates: first, the electrostatic environment which reorganizes to stabilize the nascent product charge distribution, and second the contraction of the hydride donor–acceptor distance (here CC) which facilitates the transfer[67] (the third coordinate is the displacement of the hydride particle, which adapts very fast to the two other slower coordinates). We now show that our results on EcDHFR and TmDHFR confirm both aspects, we determine their respective importances, and elucidate their change with temperature.

First, our results for both homologs reveal a strong increase in the reaction free energy barrier $\Delta G^{\ddagger}$ with increasing CC distance between the NADPH donor and folate acceptor carbon atoms in the reactant configuration (Fig. 3a). This does not mean that the CC distance is necessarily a good reaction coordinate or the only descriptor of DHFR reactivity, but there is a clear correlation between this distance and the free energy barrier. Reactant configurations with longer CC distances thus require a greater contraction free energy cost to reach short transition-state (TS) donor-acceptor separations, where the hydride can be more easily transferred. We note that the same CC distance was found at the TS (Supplementary Figure S5), independently of the starting reactant configuration; in agreement with prior simulations[44] this TS CC distance is short, which implies that tunneling is limited, in accordance with the small experimental kinetic isotope effect values for both homologs.[68] Our results further reveal that i) the variation in $\Delta G^{\ddagger}$ induced by typical CC fluctuations is much larger than that caused by an increase in temperature for a fixed CC distance (see Supplementary Tables S3 and S4); ii) for a given CC distance in the open conformation, $\Delta G^{\ddagger}$ is approximately equal in Ec and TmDHFR (see Fig. 3a). This leads to two major and general conclusions that i) the difference in sequences does not affect the reaction barrier directly, but rather indirectly by a change in the conformational distributions, and ii) the difference

in apparent activation energies is caused by the different ways in which the conformational distributions change with temperature. We now continue with the strong impact of the Met20 loop configuration on the CC distance, which is shown in Fig. 3b. Closed conformations always imply short CC distances, while open conformations lead to broad CC distributions with longer average CC distances. Figures 3c-d show that the loop opening above 330 K in EcDHFR leads to a sudden broadening of the CC distribution and an increase in the average CC distance, while for TmDHFR, the distribution remains broad and extends to large distances at all temperatures. As recently shown, the open loop conformation in TmDHFR is not a consequence of dimerization, but rather stems from its different sequence.[36,37] Indeed, the Met20 loop remains open in the hypothetical, isolated monomer.[37]

Another important role played by the Met20 loop is to change the electrostatic environment next to the transferred hydride. The apolar environment provided by the closed Met20 loop facilitates the disappearance of the positive charge on the N5 atom during the hydride transfer step (see Fig. 1b), while in the open configuration, this site is exposed to the polar aqueous solvent, which increases the reaction free energy barrier[67,69] (Supplementary Section III.2 and Figures S6 and S7). This explains why for the same short CC distances in EcDHFR (2.98–3.33 $\mathring{A}$), $\Delta G^{\ddagger}$ is much smaller when the loop is closed ($\langle \Delta G^{\ddagger} \rangle_c$=13.7 kcal/mol) than when it is open ($\Delta G^{\ddagger} = 15$–22 kcal/mol), as shown in Fig. 3b. The loop conformation thus has a major effect on both the CC distribution (which itself influences the free energy barrier), and the free energy barrier for very short CC distances.

Model for activity temperature dependence and optimum temperature

We now combine our results on the relationship between reaction free energy barrier and enzyme conformation and on the temperature-induced structural changes, in order to determine how the rate

constant changes with temperature for different homologs. We specifically focus on the molecular origin of the different activation energies in the two homologs, and on the key features that determine the optimum catalytic activity temperature.

Our results show that the differences between the two homologs regarding their catalytic activities and their temperature dependences are caused by their different conformations, and by the great sensitivity of the reaction free energy barrier to the Met20 loop conformation, which affects both the local electrostatic environment and the CC distance distribution. We now use this molecular picture to develop a model describing the temperature dependence of enzyme activity, and identify which features determine the apparent activation energy and the optimum activity temperature.

Our goal is to determine the experimentally accessible average rate constant, which results from a distribution of conformations with different activities. We first estimate the average rate constant of folded (and thus active) protein conformations $\langle k_f(T)\rangle$. Using the results in Fig. 3, $\langle k_f(T)\rangle$ is determined by the fraction of closed loop configurations with a $\langle \Delta G^{\ddagger}\rangle_c$ barrier, and the complementary fraction of open loop configurations where the free energy barrier $\Delta G^{\ddagger}_{open}$ increases with the CC distance. Assuming a fast conformational equilibrium as compared to the reaction timescale (see Supplementary Information Section IV.4), a Transition-State Theory (TST) description thus leads to

$$\langle k_f(T)\rangle = \frac{k_B T}{h}\Big[P_c(T)e^{-\langle \Delta G^{\ddagger}\rangle_c/k_B T} + (1-P_c(T))\int_0^{\infty} \mathrm{d}CC\; p_{CC}(T) e^{-\frac{\Delta G^{\ddagger}_{open}(CC)}{k_B T}}\Big], \quad (1)$$

where P_c is the fraction of folded proteins in the closed state, and p_{CC} is the probability distribution of CC distances in the (folded) open state. While prior studies[36,44] have complemented this TST rate constant with corrections due to transmission factor, tunneling and zero point energies, these are not expected to significantly change with temperature (as shown by the experimental kinetic isotope effect[68]) and are not considered here. We note the importance of averaging over the rate $\langle k \rangle$ rather than considering the rate corresponding to the average free energy barrier $\langle \Delta G^{\ddagger} \rangle$ (see Supplementary Section IV). The small fraction of low $\Delta G^{\ddagger}$ conformations thus bring a very important contribution to $\langle k \rangle$, and a direct average of the free-barrier overall the conformational ensemble would clearly lead to an incorrect picture of its temperature dependence. At the other end of the conformational ensemble, the fraction of conformations with very high free energy barriers, corresponding to large CC distances, bring a negligible contribution.

The $\langle k_f(T) \rangle$ values for both homologs are plotted in Fig. 4a. Our simulation results suggest that that the folded-state rate constant can be well approximated by an Arrhenius temperature dependence

$$\langle k_f(T) \rangle \simeq A\, e^{-E_a/k_B T}\,, \tag{2}$$

where E_a is the reaction activation energy and A is a temperature-independent prefactor (to facilitate the comparison of simulated and experimental rate temperature dependences, in the following A will be adjusted to reproduce the ambient temperature experimental rate constant). The activation energy E_a (reported in Table 1) significantly differs from the free energy barriers since it includes contributions from temperature-induced conformational changes within the folded state, as discussed in detail further below. The approximation of a constant E_a is supported for the present systems by their linear Arrhenius plots far from T_m in the simulations.

We can then determine the overall rate constant $\langle k(T) \rangle$ which is measured experimentally, by reweighting $\langle k_f(T) \rangle$ with the folded protein fraction $P_f(T)$,

$$\langle k(T) \rangle = P_f(T) \langle k_f(T) \rangle \, . \tag{3}$$

Here we use a simplified two-state picture where unfolded proteins are inactive, which echoes concepts used in previous approaches[16]. The uncatalyzed reaction rate constant is assumed to be negligible, given the many orders of magnitude typically reported between the catalyzed and uncatalyzed rates[70]. The folded protein fraction P_f is determined by the protein stability curve obtained from our simulations,

$$P_f(T) = \left[1 + e^{-\Delta G_u / k_B T} \right]^{-1} \, , \tag{4}$$

where the unfolding free energy ΔG_u is determined by the Gibbs-Helmholtz equation (see Supplementary Section V) using the thermal stability data from the same set of simulations.[37]

The average rate constants $\langle k(T) \rangle$ predicted by our model equation (3) with the parameters obtained from our simulations are shown in Fig. 4b. To focus on the comparison between experimental and simulated rate temperature dependence, temperature axes in these Arrhenius plots are shifted to have a common origin at the melting temperature. The slope and optimum activity temperature from our model are in very good agreement with both values obtained directly from our simulations and experiments at a series of temperatures. We therefore use this model to first analyze the behavior far below the melting temperature to explain the apparent activation energy, and then focus on the vicinity of the melting temperature to identify what governs the optimum

activity temperature.

Activation energy At temperatures far below the activity optimum, the behavior of EcDHFR and TmDHFR is Arrhenian (Fig. 4b) and the slope of $\ln\langle k(T)\rangle$ with $1/T$ provides the reaction activation energy E_a. This is consistent with the Arrhenius dependence of $\langle k_f(T)\rangle$ (Fig. 4a) far from melting, where all enzymes are folded. This activation energy is found to be much larger in TmDHFR than in EcDHFR (see Table 1), in agreement with experimental data.[29,30] More generally, this is in accord with the trend observed among psychrophilic, mesophilic and thermophilic enzymes, that the rate constant acceleration with increasing temperature is more pronounced for enzymes whose optimum activity temperature is higher.[71]

However, while the experimentally accessible E_a is often used as an estimate of the intrinsic reaction free energy barrier, our results show that E_a is strikingly smaller than the free energy barrier (Table 1). In the traditional picture, E_a is usually assimilated to the $\Delta H^\ddagger$ enthalpic component and this difference would be interpreted as arising from a large $\Delta S^\ddagger$ entropy increase in the transition state (a difference in activation entropies was recently suggested to explain the different E_a in mesophilic and cold-adapted enzymes).[19] However, as detailed in the Supplementary Section III.3, our results suggest that in a given conformation (i.e., fixed CC distance), $\Delta S^\ddagger$ is very small. In contrast, our results reveal that E_a includes an important contribution from the temperature-induced changes in the conformational equilibrium, and in particular in the fraction of closed-loop conformations and in the CC donor-acceptor distribution. The reaction enthalpic barrier sensitively depends on the cost to contract the CC distance, and a temperature increase causes a change in the conformational distribution, and thus in the average reaction barrier. Therefore, the large difference between E_a and the free energy barrier does not come from an entropic contribution to the chemical barrier and instead arises from temperature-induced changes in the protein conformational distribution. These changes are most pronounced in EcDHFR due to the Met20 loop opening

at high temperature, which increases $\Delta G^{\ddagger}$ and lowers the apparent E_a; in contrast in TmDHFR, these changes are more limited and E_a is closer to the free energy barrier.

We pause to comment on the effect of dimerization on the calculated rate constants for TmDHFR (Eq. 3). First, the dimeric nature of this enzyme does not affect the applicability of our model. Indeed, its key ingredients are (i) the free-energy barriers for given active site conformations (which are identical in EcDHFR and TmDHFR, and thus do not significantly change with sequence or oligomerization); and (ii), the temperature dependences of enzyme conformations, which can be determined regardless of the enzyme oligomerization state. Finally, we note that while dimerization strongly enhances the TmDHFR thermal stability (see e.g. refs.[36,37]), experiments[62] that compared the native TmDHFR dimer with a monomeric mutant concluded that both enzymes display similar absolute rate constants at a given temperature and similar rate temperature dependences.

We stress that the important role of temperature-induced conformational shifts revealed here contrasts with the traditional picture focusing on conformational flexibility, i.e. on the width of the conformational distribution. If we neglect the effect of a distribution of conformations in Eq. 1 by taking $P_c = 1$ at all temperatures, we obtain a value of $E_a = 13.0$ kcal/mol for EcDHFR, which is in contradiction with the experimental values (3.7–7.1 kcal/mol[28–30]). This clearly justifies both our enhanced sampling strategy and the major impact of the conformation ensemble on the catalytic rate temperature dependence. Our results thus show that imposing conformational restraints on the mesophilic enzyme would induce a thermophilic behavior. The same behavior was observed in a recent simulation study comparing psychrophile and mesophile.[19] While the latter work focused on rigidity and the entropy of surface groups, our model shows the major role played by conformational rigidity in general, and specifically by temperature-induced conformational shifts that affect the reaction free energy barrier.

Optimum temperature We now use our model to determine the molecular factors which govern the optimum activity temperature $T_{\rm opt}$. For both homologs, experiments show that the optimal activity is reached before melting occurs at T_m. However, the difference between $T_{\rm opt}$ and T_m differs greatly in the two systems, as shown by the experimental temperature shifts of 16 K in EcDHFR, and only 3 K in TmDHFR (see Table 1).

As shown in Figure 4b and Table 1, our model can reproduce these temperature shifts quantitatively. In equation (3), the optimum occurs when the increase in the average folded state rate constant $\langle k_f(T)\rangle$ due to the increasing thermal energy is compensated by the decreasing fraction of folded proteins $P_f(T)$. Within our model, $T_{\rm opt}$ can be shown to be approximately (see Methods)

$$\frac{T_{\rm opt}}{T_m} \simeq 1 - \frac{k_B T_m}{\Delta H_u} \ln\left[\frac{\Delta H_u}{E_a} - 1\right] . \tag{5}$$

This important result shows that the key quantity determining the temperature shift between T_m and $T_{\rm opt}$ is the ratio between the apparent activation energy E_a and the unfolding enthalpy ΔH_u. Three main regimes are thus expected (see Fig. 4c). First, when E_a is much smaller than ΔH_u (as found for EcDHFR) $\Delta H_u/E_a - 1 > 1$ and the second right-hand term in equation (5) is negative, leading to an activity optimum at a temperature significantly below T_m. We pause to note that our model could thus offer an explanation to convex Arrhenius plots observed for some enzymatic systems[7,12] and often interpreted in terms of an activation heat capacity $\Delta C_p^{\ddagger}$. Whether the convexity actually stems from the enzyme denaturation has been recently debated.[7,11,12] Our results suggest that a convexity can be observed well below T_m in Arrhenius plots solely due the interplay between the enzyme activation energy and its unfolding enthalpy. In fact, fitting the experimental hydride-transfer rates for EcDHFR[30] between 288 and 318 K, with a transition state theory expression that includes a temperature dependence of the reaction enthalpy and entropy,[12] and a

temperature reference at 298 K, leads to $\Delta C_p^{\ddagger} = -0.7$ kcal/mol/K. This is on the same order as values reported before for other enzymes.[11,12] Current explanations for non-zero $\Delta C_p^{\ddagger}$ values involve a two-state model with temperature-dependent reactant conformations bearing different reactivities,[14] and proceeding through the same transition state.[7] Our model shares some similarities with these previous suggestions, but it provides a direct and quantitative picture of the effect of multiple reactive conformations without assuming 2 or several states. Temperature-induced conformational changes directly affect E_a, whose interplay with ΔH_u determines the shift between the optimum and melting temperatures, and thus the apparent convexity of the Arrhenius plot sometimes well below T_m.

We now discuss the two other regimes. When E_a is close to $\Delta H_u/2$ (as is the case for TmDHFR), the second right-hand term in equation (5) vanishes, leading to $T_{\rm opt} \approx T_m$, as observed experimentally for TmDHFR. Finally, the third case occurs for larger E_a values ($E_a > \Delta H_u/2$), where the rate increase with temperature more than compensates the diminishing fraction of active protein conformations when approaching melting, and $T_{\rm opt} > T_m$. This situation is less common, since viable enzymes require $\Delta H_u > E_a$ in order not to unfold before the reaction is catalyzed, but cases for which $T_{\rm opt} > T_m$ have already been reported.[72,73]

Our general model applies to biocatalysts beyond the case of Ec and TmDHFR enzymes. A natural extension is the DHFR from the psychrophile *Moritella profunda*. Its melting temperature is located around 313 K (depending whether it is in the apo or holo state[74]). While hydride-transfer rates have not been measured above 303 K[75], they seem to start to plateau at this temperature (Figure S7 of ref [[75]]), suggesting that the optimal temperature is very close to this value. Within the framework of our model, the large $\approx$ 10 K difference between the optimum and melting temperature is consistent with the very low activation energy for this enzyme,[29] that is on the order of 5 kcal/mol. However, some care should be taken when interpreting these observations because

the conformational motions that occur during the catalytic cycle may be different between the two enzymes, as suggested by experiments.[75]

While the relevant ΔH_u is that of the active site structure which might be difficult to determine for proteins significantly larger than DHFR (for which our calculations showed a quantitative agreement), our model provides guidelines to determine how the optimum activity temperature changes with key enzymatic structural stability and reactivity features. We now provide two examples.

First, equation (5) provides a molecular explanation to the $T_m - T_{\rm opt}$ shifts measured for an extensive set of 100 adenylate kinase variants.[73] Our model offers a molecular picture identifying the molecular origin of the very different temperature dependences observed in homologs and of three main specific features: i) it explains the observation that T_m and $T_{\rm opt}$ shift in the same direction when comparing variants, ii) it shows why $T_{\rm opt}$ is below T_m in most cases, and iii) it explains why some systems exhibit a $T_{\rm opt}$ above T_m (Fig. 4c). This also confirms that the three above-described regimes can occur. However, in contrast to a suggestion in ref.[73], our model shows that $T_{\rm opt}$ cannot be used as a proxy to probe protein function, since $T_{\rm opt}$ is sensitive to E_a, which can significantly differ from the $\Delta H^{\ddagger}$ barrier in one configuration, due to temperature-induced conformational rearrangements.

Second, our model provides novel insight in thermal adaptation of extremophilic enzymes and shows that their optimum activity temperature is not exclusively due to either a shift in structural melting temperature or a change in reaction activation enthalpy. Experimentally, it was found[9] that, for a series of α-amylases ranging from psychrophilic to thermophilic homologs, $T_{\rm opt}$ varies much more than T_m (43 K and 32 K respectively), thus showing that T_{opt} does not uniformly shift with T_m. Our model shows that the enhanced $T_{\rm opt}$ thermal adaptation arises from a change in the reaction E_a which amplifies the effect of the T_m structural stability shift. While shifts in pro-

tein structural stability are typically[76] described by a combination of three types of stability curve changes – right-shift of the maximum stability temperature, upshift of the unfolding free energy at this temperature and broadening of the stability curve – we show (see Supplementary Section V) that for all scenarios, $T_m - T_{\rm opt}$ remains mostly determined by the reaction E_a (Supplementary Figure S12). We stress that the E_a differences found among homologs are not caused by different catalytic mechanisms, as shown by the very small change in $\Delta G^{\ddagger}$ (among extremophilic α-amylases[9], $\Delta\Delta G^{\ddagger}$=1.3 kcal/mol at 288 K while $\Delta\Delta E_a$=14 kcal/mol). In contrast, as shown for DHFR, E_a includes an important contribution from temperature-induced conformational changes. Here, it is not the overall flexibility which matters, but rather the conformational heterogeneity along key coordinates which affect the reaction barrier, e.g. the CC distance in DHFR. A greater temperature-sensitivity of this critical conformational coordinate leads to a larger decrease in E_a with respect to the reaction enthalpic barrier. Based on our DHFR results, we can thus suggest that enzymes with lower optimum temperature (as compared to their melting temperature) exhibit larger temperature-induced changes which affect the reaction barrier.

Concluding remarks

Thermal adaptation of enzymes is much more complex than a uniform shift in the melting and optimum activity temperatures or a change in reaction activation enthalpy. For example, an unexpected temperature dependence of the catalytic rate leading to convex Arrhenius plots have been observed for several systems,[11,12] and their origin has recently received a lot of attention.[7] These have been interpreted in terms of multiple conformations with different reactivities whose populations could change with temperature,[14–16] although a specific quantification of this effect at the molecular level in a real protein system remained elusive.

Here, we have studied the catalytic activity temperature dependence of DHFR and its changes

between homologs (*E. coli* and *T. maritima*) thriving at different temperatures. Experimental studies have shown that TmDHFR is less active than EcDHFR, which has been intensively studied and rationalized in previous work, and which was attributed to the different conformation of the Met20 loop that is crucial to efficient catalysis.[32–36] We have instead focused on other and yet unexplained observations. First, well below melting, Ec and TmDHFR exhibit Arrhenian behaviors, but the activation energy is much smaller in Ec than in TmDHFR. Second, the temperature of maximum activity for TmDHFR is located very close to its melting point, while it is significantly below for EcDHFR. We combined both an enhanced sampling of the enzymatic conformational space and an explicit treatment of electronic rearrangements to account for the enzyme chemical step. We show that such a combination is essential to properly describe the temperature effects on the broad distribution of conformations with different catalytic barriers. Our strategy is shown to successfully account for the large difference between the temperature dependence of the catalytic rate in both homologs. We demonstrate that neglecting a proper exploration of the conformational at increasing temperatures, which can only be achieved with enhanced sampling to overcome the large free energy barriers separating relevant states for reactivity, would fail to properly reproduce the experimental observations.

Our approach further allows to understand the molecular origin of the previously unexplained differences in the temperature dependence of catalytic rate in both homologs. In EcDHFR, a key conformational change around the active site, involving the opening of the Met20 loop, occurs upon heating at temperatures well below global melting, leading to an unusually-small activation energy as compared to TmDHFR, which in turns results in an optimum temperature that is significantly lower than its melting temperature. We develop an analytic model that successfully predicts the optimum catalytic activity temperature, and shows how it is determined by the protein unfolding equilibrium and the catalyzed reaction activation energy. Our results provide molecular grounds to the idea that an equilibrium between conformations with different reactivities can give rise

to convex Arrhenius plots.[14–16] In our picture, the optimal temperature and the activation heat capacity (that can be determined ad hoc from the temperature dependence of the catalytic rate[11,12]), are due to the interplay between the reaction activation energy (which results in part from the presence of multiple conformations with different reactivities) and thermal denaturation. This general model is shown to be valid for a variety of other enzyme systems, and suggests that enzyme thermal adaptation results from both structural stability curve shifts and temperature effects on the distribution of conformations with different reaction free energy barriers.

As a final conclusion, we stress that the temperature dependence of the enzyme catalytic rate through an apparent Arrhenius activation energy does not only result from the intrinsic chemistry of the catalyzed reaction, but it also depends on temperature-induced conformational changes. In turn, the onset of the decline in activity, which can occur well below the enzyme melting temperature, is determined by the activation energy and by how it compares to the enzyme denaturation enthalpy. Our model predicts that in general, small activation energies would result in a large shift between the optimum and melting temperatures, and thus to apparent convex Arrhenius plots at temperatures well below melting. Larger activation energies would instead result in optimum temperatures closer to melting. These results provide guidelines for the design of biocatalysts with tailored thermal properties.

Methods

Extensive details about system preparation, employed force-fields, and simulation methodology are provided in the Supplementary Information. A brief summary of the key employed simulation strategies and definitions is given below.

Protein conformational sampling To sample relevant protein conformations at several temperatures, we ran solute-tempering Hamiltonian replica exchange (REST2[55,56]) simulations with an

all-atom description of the protein and its solvent, as described in detail in our recent work[37].

Protein reactivity. The electronic structure of the substrate and cofactor is modeled as a linear combination of two coupled resonance states, each described by a classical force-field. Here we use a set of parameters that were shown to successfully describe the reaction in EcDHFR at ambient temperature[45]. $\Delta G^{\ddagger}$ is determined as the free energy barrier along the energy gap between these two diabatic states (see Supplementary Figure S1). In order to account for the broad conformational distribution and its impact on the reaction rate, 10 independent $\Delta G^{\ddagger}$ calculations are performed for each homolog and at each temperature below T_m, from uncorrelated initial configurations taken from the REST2 sampling. These configurations were extracted from our REST2 trajectories (rescaled potential energy but ambient temperature) but then propagated at the corresponding physical temperature (with no more potential energy rescaling).

Met20 loop conformation. We discriminate between the Met20 loop open and closed conformations by analyzing the distance between this loop and the C helix (loop-helix distance). In particular, we look at the distance between the EcDHFR Asn18 and His45 C_α atoms, which, after structural alignment, corresponds to the TmDHFR Val19 and Ile46 C_α atoms. The loop is considered closed if this distance is between 6 and 8 Å (see Supplementary Section II.5.).

Derivation of Equation (5). T_{opt} is defined by

$$\left.\frac{d\langle k(T)\rangle}{dT}\right|_{T=T_{opt}} = 0, \tag{6}$$

or equivalently

$$\left.\frac{d\ln\langle k(T)\rangle}{dT}\right|_{T=T_{opt}} = 0. \tag{7}$$

So, by combining it with Eqs. 2,3, and 4, $T_{\rm opt}$ is the solution of

$$k_B T^2 \frac{d\ln\langle k(T)\rangle}{dT} = \left[1 - \frac{1}{1+e^{-\Delta G_u/k_B T}}\right]\left(T\frac{d\Delta G_u}{dT} - \Delta G_u\right) + E_a = 0, \tag{8}$$

where, using the Gibbs-Helmholtz equation (see Supplementary Section V) for the determination of ΔG_u ,

$$\frac{d\Delta G_u}{dT} = -\frac{\Delta H_u}{T_m} - \Delta C_p \ln\left(T/T_m\right). \tag{9}$$

T_{opt} is therefore the solution of

$$\left[1 - \frac{1}{1+e^{-\Delta G_u/k_B T}}\right]\left[\Delta C_p\left(T_m - T_{opt}\right) - \Delta H_u\right] + E_a = 0. \tag{10}$$

A first approach is to solve equation (10) numerically. A second approach is to consider that T_{opt} is close enough to T_m to perform a first-order expansion of equation (10) in $\epsilon = 1 - T_{opt}/T_m$, and solve for T_{opt}. The first step leads to

$$e^{-\Delta G_u(\epsilon)/(k_B T_m(1-\epsilon))} = \frac{E_a}{\Delta H_u - E_a - \Delta C_p T_m \epsilon} \tag{11}$$

whose first-order ϵ expansion yields

$$-\ln\left[\frac{\Delta H_u - E_a}{E_a}\right] + \epsilon\left[\frac{T_m \Delta C_p}{E_a - \Delta H_u} + \frac{\Delta H_u}{k_B T_m}\right] \simeq 0, \tag{12}$$

leading to the $T_{\rm opt}$ approximate solution

$$\frac{T_{opt}}{T_m} \simeq 1 - \frac{1}{\frac{\Delta H_u}{k_B T_m} + \frac{T_m \Delta C_p}{\Delta H_u - E_a}} \ln\left[\frac{\Delta H_u}{E_a} - 1\right]. \tag{13}$$

Given the typical values of $\Delta H_u, \Delta C_p, T_m, E_a$ found for EcDHFR and TmDHFR, this can be simplified into

$$\frac{T_{opt}}{T_m} \simeq 1 - \frac{k_B T_m}{\Delta H_u} \ln\left(\frac{\Delta H_u}{E_a} - 1\right). \tag{14}$$

Comparing the numerical solution of equation (10) and the approximate result equation (14) for EcDHFR and TmDHFR shows that the two temperatures differ by less than 1 K.

[1] Clarke, A. The Thermal Limits to Life on Earth. *Int. J. Astrobiol.* **13**, 141–154 (2014).

[2] Somero, G. N. Temperature Adaptation of Enzymes: Biological Optimization Through Structure-Function Compromises. *Annu. Rev. Ecol. Syst.* **9**, 1–29 (1978).

[3] Vieille, C. & Zeikus, G. J. Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. *Microbiol. Mol. Biol. Rev.* **65**, 1–43 (2001).

[4] Feller, G. Protein Stability and Enzyme Activity at Extreme Biological Temperatures. *J. Phys. Condens. Matter* **22**, 323101 (2010).

[5] Somero, G. N. Proteins and Temperature. *Annu. Rev. Physiol.* **57**, 43–68 (1995).

[6] Arnold, F. H., Wintrode, P. L., Miyazaki, K. & Gershenson, A. How Enzymes Adapt: Lessons from Directed Evolution. *Trends Biochem. Sci.* **26**, 100–106 (2001).

[7] Arcus, V. L. & Mulholland, A. J. Temperature, Dynamics, and Enzyme-Catalyzed Reaction Rates. *Annu. Rev. Biophys.* **49**, 163–180 (2020).

[8] DeLong, J. P. *et al.* The Combined Effects of Reactant Kinetics and Enzyme Stability Explain the Temperature Dependence of Metabolic Rates. *Ecol. Evol.* **7**, 3940–3950 (2017).

[9] Cipolla, A., Delbrassine, F., Da Lage, J.-L. & Feller, G. Temperature Adaptations in Psychrophilic, Mesophilic and Thermophilic Chloride-Dependent Alpha-Amylases. *Biochimie* **94**, 1943–1950 (2012).

[10] Georlette, D. *et al.* Structural and Functional Adaptations to Extreme Temperatures in Psychrophilic, Mesophilic, and Thermophilic DNA Ligases. *J. Biol. Chem.* **278**, 37015–37023 (2003).

[11] Nguyen, V. *et al.* Evolutionary Drivers of Thermoadaptation in Enzyme Catalysis. *Science* **355**, 289–294 (2017).

[12] Arcus, V. L. *et al.* On the Temperature Dependence of Enzyme-Catalyzed Rates. *Biochemistry* **55**, 1681–1688 (2016).

[13] Almeida, V. M. & Marana, S. R. Optimum Temperature May Be a Misleading Parameter in Enzyme Characterization and Application. *PLoS One* **14**, e0212977 (2019).

[14] Daniel, R. M. & Danson, M. J. A New Understanding of How Temperature Affects the Catalytic Activity of Enzymes. *Trends Biochem. Sci.* **35**, 584–591 (2010).

[15] Kavanau, J. L. Enzyme Kinetics and the Rate of Biological Processes. *J. Gen. Physiol.* **34**, 193–209 (1950).

[16] Truhlar, D. & Kohen, A. Convex Arrhenius Plots and Their Interpretation. *Proc. Natl. Acad. Sci. USA* **98**, 848–851 (2001).

[17] Low, P. S., Bada, J. L. & Somero, G. N. Temperature Adaptation of Enzymes: Roles of the Free Energy, the Enthalpy, and the Entropy of Activation. *Proc. Natl. Acad. Sci. USA* **70**, 430–2 (1973).

[18] Wolfenden, R. Primordial Chemistry and Enzyme Evolution in a Hot Environment. *Cell. Mol. Life Sci.* **71**, 2909–2915 (2014).

[19] Isaksen, G. V., Åqvist, J. & Brandsdal, B. O. Enzyme Surface Rigidity Tunes the Temperature Dependence of Catalytic Rates. *Proc. Natl. Acad. Sci.* **113**, 7822–7827 (2016).

[20] Sočan, J., Kazemi, M., Isaksen, G. V., Brandsdal, B. O. & Åqvist, J. Catalytic adaptation of psychrophilic elastase. *Biochemistry* **57**, 2984–2993 (2018).

[21] Masgrau, L. & Truhlar, D. G. The Importance of Ensemble Averaging in Enzyme Kinetics. *Acc. Chem. Res.* **48**, 431–438 (2015).

[22] Stirnemann, G. & Sterpone, F. Mechanics of Protein Adaptation to High Temperatures. *J. Phys. Chem. Lett.* **8**, 5884–5890 (2017).

[23] Zeiske, T., Stafford, K. A. & Palmer, A. G. Thermostability of Enzymes from Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **12**, 2489–2492 (2016).

[24] Garcia-Viloca, M., Gao, J., Karplus, M. & Truhlar, D. G. How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations. *Science* **303**, 186–195 (2004).

[25] Warshel, A. *et al.* Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **106**, 3210–3235 (2006).

[26] Benkovic, S. J. & Hammes-Schiffer, S. A Perspective on Enzyme Catalysis. *Science* **301**, 1196–1202 (2003).

[27] Schnell, J. R., Dyson, H. J. & Wright, P. E. Structure, Dynamics, and Catalytic Function of Dihydrofolate Reductase. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 119–140 (2004).

[28] Sikorski, R. S. *et al.* Tunneling and Coupled Motion in the Escherichia Coli Dihydrofolate Reductase Catalysis. *J. Am. Chem. Soc.* **126**, 4778–4779 (2004).

[29] Loveridge, E. J. *et al.* the Role of Large-Scale Motions in Catalysis by Dihydrofolate Reductase. *J. Am. Chem. Soc.* **133**, 20561–20570 (2011).

[30] Maglia, G., Javed, M. H. & Allemann, R. K. Hydride Transfer During Catalysis by Dihydrofolate Reductase from Thermotoga Maritima. *Biochem. J.* **374**, 529–535 (2003).

[31] Guo, J., Luk, L. Y. P., Loveridge, E. J. & Allemann, R. K. Thermal Adaptation of Dihydrofolate Reductase from the Moderate Thermophile Geobacillus Stearothermophilus. *Biochemistry* **53**, 2855–2863 (2014).

[32] Dams, T. & Jaenicke, R. Stability and Folding of Dihydrofolate Reductase from the Hyperthermophilic Bacterium Thermotoga Maritima. *Biochemistry* **38**, 9169–78 (1999).

[33] Dams, T. *et al.* Homo-Dimeric Recombinant Dihydrofolate Reductase from Thermotoga Maritima Shows Extreme Intrinsic Stability. *Biol. Chem.* **379**, 367–71 (1998).

[34] Dams, T. *et al.* The Crystal Structure of Dihydrofolate Reductase from Thermotoga Maritima: Molecular Features of Thermostability. *J. Mol. Biol.* **297**, 659–672 (2000).

[35] Pang, J. & Allemann, R. K. Molecular Dynamics Simulation of Thermal Unfolding of Thermatoga Maritima DHFR. *Phys. Chem. Chem. Phys.* **9**, 711–718 (2007).

[36] Ruiz-Pernía, J. J., Tuñón, I., Moliner, V. & Allemann, R. K. Why Are Some Enzymes Dimers? Flexibility and Catalysis in Thermotoga Maritima Dihydrofolate Reductase. *ACS Cat.* **9**, 5902–5911 (2019).

[37] Maffucci, I., Laage, D., Stirnemann, G. & Sterpone, F. Differences in thermal structural changes and melting between mesophilic and thermophilic dihydrofolate reductase enzymes. *submitted* 10.26434/chemrxiv.12369737 (2020).

[38] Radkiewicz, J. L. & Brooks, C. L. Protein Dynamics in Enzymatic Catalysis: Exploration of Dihydrofolate Reductase. *J. Am. Chem. Soc.* **122**, 225–231 (2000).

[39] Agarwal, P. K., Billeter, S. R., Rajagopalan, P. T. R., Benkovic, S. J. & Hammes-Schiffer, S. Network of Coupled Promoting Motions in Enzyme Catalysis. *Proc. Natl. Acad. Sci.* **99**, 2794–2799 (2002).

[40] Sawaya, M. R. & Kraut, J. Loop and Subdomain Movements in the Mechanism of Escherichia Coli Dihydrofolate Reductase: Crystallographic Evidence. *Biochemistry* **36**, 586–603 (1997).

[41] Åqvist, J. & Warshel, A. Simulation of Enzyme Reactions Using Valence Bond Force Fields and Other Hybrid Quantum/classical Approaches. *Chem. Rev.* **93**, 2523–2544 (1993).

[42] Warshel, A., Sussman, F. & Hwang, J.-K. Evaluation of Catalytic Free Energies in Genetically Modified Proteins. *J. Mol. Biol.* **201**, 139–159 (1988).

[43] Kamerlin, S. C. L. & Warshel, A. The EVB As a Quantitative Tool for Formulating Simulations and Analyzing Biological and Chemical Reactions. *Faraday Discuss.* **145**, 71–106 (2010).

[44] Agarwal, P. K., Billeter, S. R. & Hammes-Schiffer, S. Nuclear Quantum Effects and Enzyme Dynamics in Dihydrofolate Reductase Catalysis. *J. Phys. Chem. B* **106**, 3283–3293 (2002).

[45] Liu, C. T. *et al.* Functional Significance of Evolving Protein Sequence in Dihydrofolate Reductase from Bacteria to Humans. *Proc. Natl. Acad. Sci.* **110**, 10159–10164 (2013).

[46] Clark, A. C. & Frieden, C. Native Escherichia Coli and Murine Dihydrofolate Reductases Contain Late-Folding Non-Native Structures. *J. Mol. Biol.* **285**, 1765–1776 (1999).

[47] Ionescu, R. M., Smith, V. F., O'Neill, J. C., Jr & Matthews, C. R. Multistate Equilibrium Unfolding of Escherichia Coli Dihydrofolate Reductase: Thermodynamic and Spectroscopic Description of the Native, Intermediate, and Unfolded Ensembles. *Biochemistry* **39**, 9540–9550 (2000).

[48] Ohmae, E., Kurumiya, T., Makino, S. & Gekko, K. Acid and Thermal Unfolding of Escherichia Coli Dihydrofolate Reductase. *J. Biochem.* **120**, 946–953 (1996).

[49] Kasper, J. R., Liu, P.-F. & Park, C. Structure of a Partially Unfolded Form of Escherichia Coli Dihydrofolate Reductase Provides Insight into Its Folding Pathway. *Protein Sci.* **23**, 1728–1737 (2014).

[50] McElheny, D., Schnell, J. R., Lansing, J. C., Dyson, H. J. & Wright, P. E. Defining the Role of Active-Site Loop Fluctuations in Dihydrofolate Reductase Catalysis. *Proc. Natl. Acad. Sci. USA* **102**, 5032–5037 (2005).

[51] Arora, K. & Brooks, C. L. Multiple Intermediates, Diverse Conformations, and Cooperative Conformational Changes Underlie the Catalytic Hydride Transfer Reaction of Dihydrofolate Reductase. *Top. Curr. Chem.* **337**, 165–187 (2013).

[52] Rocco, A. G. *et al.* Characterization of the Protein Unfolding Processes Induced by Urea and Temperature. *Biophys. J.* **94**, 2241–2251 (2008).

[53] Stirnemann, G., Kang, S.-g., Zhou, R. & Berne, B. J. How Force Unfolding Differs from Chemical Denaturation. *Proc. Natl. Acad. Sci. USA* **111**, 3413–3418 (2014).

[54] Wang, T. & Wade, R. C. On the Use of Elevated Temperature in Simulations To Study Protein Unfolding Mechanisms. *J. Chem. Theory Comput.* **3**, 1476–1483 (2007).

[55] Wang, L., Friesner, R. A. & Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B* **115**, 9431–9438 (2011).

[56] Stirnemann, G. & Sterpone, F. Recovering Protein Thermal Stability Using All-Atom Hamiltonian Replica-Exchange Simulations in Explicit Solvent. *J. Chem. Theory Comput.* **11**, 5573–5577 (2015).

[57] Katava, M. *et al.* Critical Structural Fluctuations of Proteins upon Thermal Unfolding Challenge the Lindemann Criterion. *Proc. Natl. Acad. Sci.* **114**, 9361–9366 (2017).

[58] Kührová, P., De Simone, A., Otyepka, M. & Best, R. B. Force-Field Dependence of Chignolin Folding and Misfolding: Comparison with Experiment and Redesign. *Biophys. J.* **102**, 1897–1906 (2012).

[59] Paschek, D., Hempel, S. & García, A. E. Computing the Stability Diagram of the Trp-Cage Miniprotein. *Proc. Natl. Acad. Sci. USA* **105**, 17754–17759 (2008).

[60] Yang, C., Jang, S. & Pak, Y. A Fully Atomistic Computer Simulation Study of Cold Denaturation of a β-Hairpin. *Nat. Commun.* **5**, 5773 (2014).

[61] Adamczak, B., Kogut, M. & Czub, J. Effect of Osmolytes on the Thermal Stability of Proteins: Replica Exchange Simulations of Trp-Cage in Urea and Betaine Solutions. *Phys. Chem. Chem. Phys.* **20**, 11174–11182 (2018).

[62] Guo, J., Loveridge, E. J., Luk, L. Y. P. & Allemann, R. K. Effect of Dimerization on Dihydrofolate Reductase Catalysis. *Biochemistry* **52**, 3881–3887 (2013).

[63] Liao, Q. *et al.* Loop Motion in Triosephosphate Isomerase Is Not a Simple Open and Shut Case. *J. Am. Chem. Soc.* **140**, 15889–15903 (2018).

[64] Mhashal, A. R., Vardi-Kilshtain, A., Kohen, A. & Major, D. T. The Role of the Met 20 Loop in the Hydride Transfer in Escherichia Coli Dihydrofolate Reductase. *J. Biol. Chem.* **292**, 14229–14239 (2017).

[65] Mhashal, A. R., Pshetitsky, Y., Cheatum, C. M., Kohen, A. & Major, D. T. Evolutionary Effects on Bound Substrate pKa in Dihydrofolate Reductase. *J. Am. Chem. Soc.* **140**, 16650–16660 (2018).

[66] Staib, A., Borgis, D. & Hynes, J. T. Proton Transfer in Hydrogen-bonded Acid Base Complexes in Polar Solvents. *J. Chem. Phys.* **102**, 2487–2505 (1995).

[67] Roca, M., Liu, H., Messer, B. & Warshel, A. On the Relationship Between Thermal Stability and Catalytic Power of Enzymes. *Biochemistry* **46**, 15076–15088 (2007).

693 [68] Loveridge, E. J., Behiry, E. M., Swanwick, R. S. & Allemann, R. K. Different Reaction
694 Mechanisms for Mesophilic and Thermophilic Dihydrofolate Reductases. *J. Am. Chem. Soc.*
695 **131**, 6926–6927 (2009).

696 [69] Luk, L. Y. P., Loveridge, E. J. & Allemann, R. K. Different Dynamical Effects in Mesophilic
697 and Hyperthermophilic Dihydrofolate Reductases. *J. Am. Chem. Soc.* **136**, 6862–6865
698 (2014).

699 [70] Wolfenden, R. & Snider, M. J. The Depth of Chemical Time and the Power of Enzymes As
700 Catalysts. *Acc. Chem. Res.* **34**, 938–945 (2001).

701 [71] Georlette, D. *et al.* Some like It Cold: Biocatalysis at Low Temperatures. *FEMS Microbiol.*
702 *Rev.* **28**, 25–42 (2004).

703 [72] Daniel, R. M. *et al.* The Molecular Basis of the Effect of Temperature on Enzyme Activity.
704 *Biochem. J.* **425**, 353–360 (2009).

705 [73] Howell, S. C., Inampudi, K. K., Bean, D. P. & Wilson, C. J. Understanding Thermal Adap-
706 tation of Enzymes Through the Multistate Rational Design and Stability Prediction of 100
707 Adenylate Kinases. *Structure* **22**, 218–229 (2014).

708 [74] Evans, R. M. *et al.* Catalysis by Dihydrofolate Reductase from the Psychropiezophile
709 Moritella Profunda. *ChemBioChem* **11**, 2010–2017 (2010).

710 [75] Behiry, E. M., Evans, R. M., Guo, J., Loveridge, E. J. & Allemann, R. K. Loop Interactions
711 During Catalysis by Dihydrofolate Reductase from Moritella Profunda. *Biochemistry* **53**,
712 4769–4774 (2014).

713 [76] Razvi, A. & Scholtz, J. M. Lessons in Stability from Thermophilic Proteins. *Protein Sci.* **15**,
714 1569–1578 (2006).

[77] Gekko, K. *et al.* Effects of Point Mutation in a Flexible Loop on the Stability and Enzymatic Function of Escherichia Coli Dihydrofolate Reductase1. *J. Biochem.* **113**, 74–80 (1993).

Acknowledgements The research leading to these results has received funding from PSL University (ANR-10-IDEX-0001-02 PSL*) through a collaborative chemistry research program, from the European Research Council (FP7/2007-2013, grant agreements 258748 to F.S. and 279977 to D.L.), and from the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01). Part of this work was performed using HPC resources from LBT, ENS, GENCI [CINES and TGCC] (Grants x201776818 and x2017077156) and PRACE allocation (Pra13-3298, 13th call).

Competing Interests The authors declare that they have no competing financial interests.

Correspondence Correspondence and requests for materials should be addressed to D.L. (email: damien.laage@ens.fr), F.S. (email: sterpone@ibpc.fr), G.S. (email: stirnemann@ibpc.fr)

Biographical Information

Guillaume Stirnemann was born in France in 1987. After undergraduate studies at the École normale supérieure (ENS) and a Master's thesis with Pablo Debenedetti at Princeton University, he received his Ph. D. (2011) from ENS and Sorbonne Université (Paris, France), under the supervision of Damien Laage. He was then a postdoctoral fellow at Columbia University with Bruce Berne and later at the Institut de Biologie Physico-Chimique (IBPC) in Paris with Fabio Sterpone. In 2014, he was recruited as a CNRS researcher at IBPC where the current interests of his group include the stability, the mechanical properties and the reactivity of biomolecules, and transport

phenomena in aqueous solutions, with a special emphasis on questions related to the origins of life.

Table 1. Thermodynamic data from our simulations (stability data from ref.[37]) and from experiments.

DHFR		T_m K	ΔH_u kcal/mol	$\Delta G^{\ddagger}_{298\mathrm{K}}$ [a] kcal/mol	E_a kcal/mol	$T_{\mathrm{m}} - T_{\mathrm{opt}}$ K
Ec	sim	396	39.2	13.3	5.6	14
	exp	326[48]	44.7[77]	14.3[29]	3.7–7.1[28–30]	16[30]
Tm	sim	424	37.0	18.0	15.8	3
	exp	356[30]		18.5[29]	12.8–18.1[29,30]	3[30]

[a] effective free-energy barrier estimated as $\Delta G^{\ddagger}_{298\mathrm{K}} \approx -k_B T_{298\mathrm{K}} \ln\left[\frac{h}{k_B T_{298\mathrm{K}}}\langle k\left(T_{298\mathrm{K}}\right)\rangle\right]$ with $\langle k\left(T_{298\mathrm{K}}\right)\rangle$ from Eq. 3

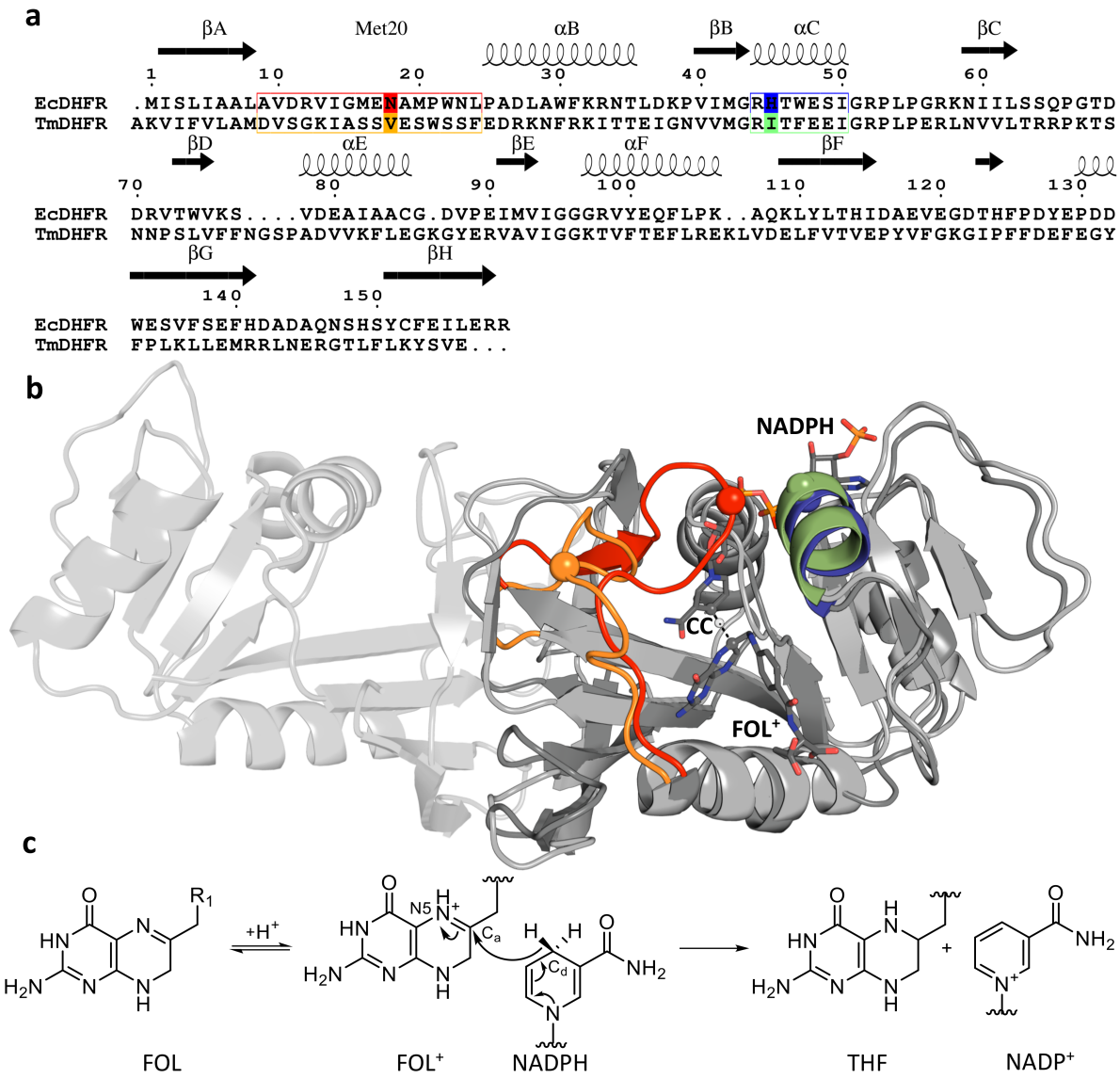

Figure 1. Ec- and TmDHFR structure and catalyzed reaction (**a** and **b**) Aligned sequences (a) and superimposed crystal structures (b) of EcDHFR and TmDHFR, highlighting the positions of the Met20 loop (red for Ec and orange for Tm), of the αC helix (blue for Ec and green for Tm). The positions of Asn18/His45 (EcDHFR) and Val19/Ile46 (TmDHFR), which are used to determine the open and closed Met20 conformations based on the Met20 loop – αC helix distance, are shown as plain boxes in (a) and balls in (b). NADPH and folate are shown in a licorice representation, and the donor carbon, acceptor carbon and transferred hydride atoms as balls. Note that the second domain in the TmDHFR dimer appears in light gray on the left hand side. (**c**) Chemical step catalyzed by DHFR: fast protonation equilibrium (not considered explicitly here) followed by the hydride transfer reaction step.

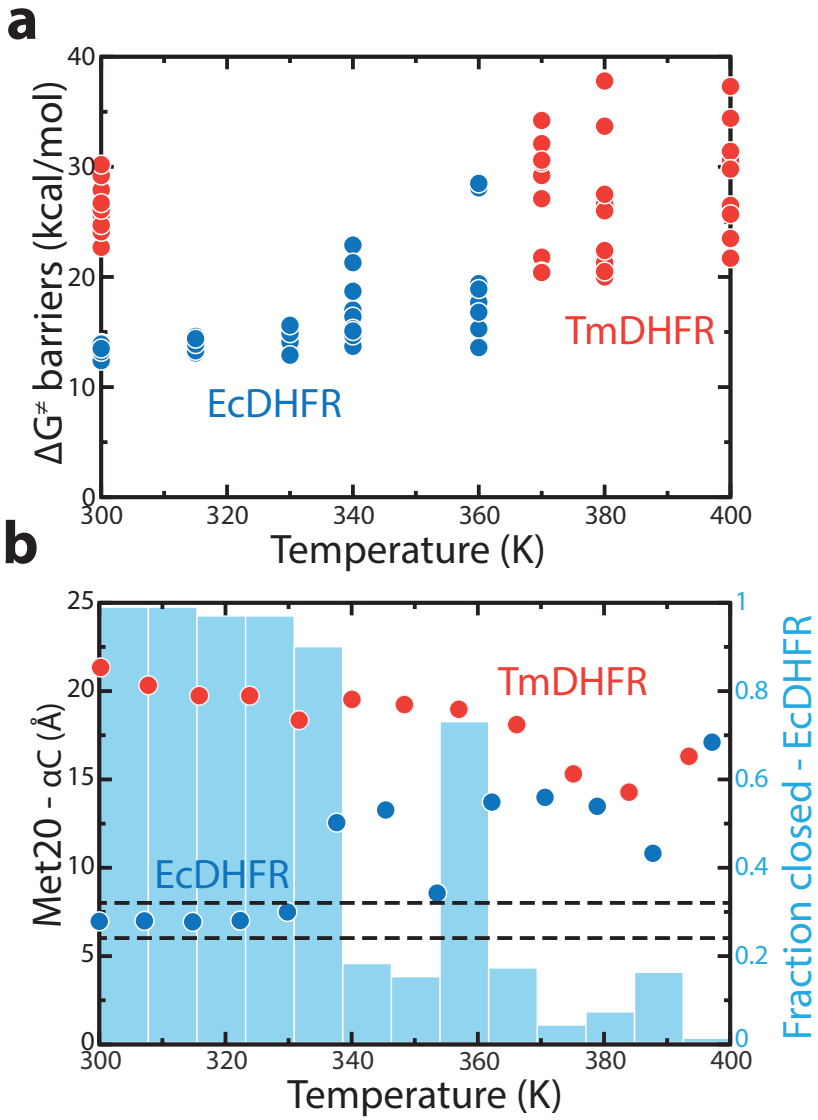

Figure 2. DHFR reactivity (**a**) Hydride transfer $\Delta G^{\ddagger}$ barrier in Ec and TmDHFR for 10 independent simulations at each temperature (Supplementary Tables S3 and S4). (**b**) Met20 loop conformations as a function of temperature estimated by following the Met20 loop – αB distance in EcDHFR (blue dots) and TmDHFR (red dots). The dots indicate the average distance at each temperature. Closed loop conformations correspond to distances between 6 and 8 Å (zone in between the black dashed lines), while open conformations are around 10–20 Å. The blue bars indicate the fraction of closed conformations for EcDHFR (for TmDHFR, the closed conformation is never found).

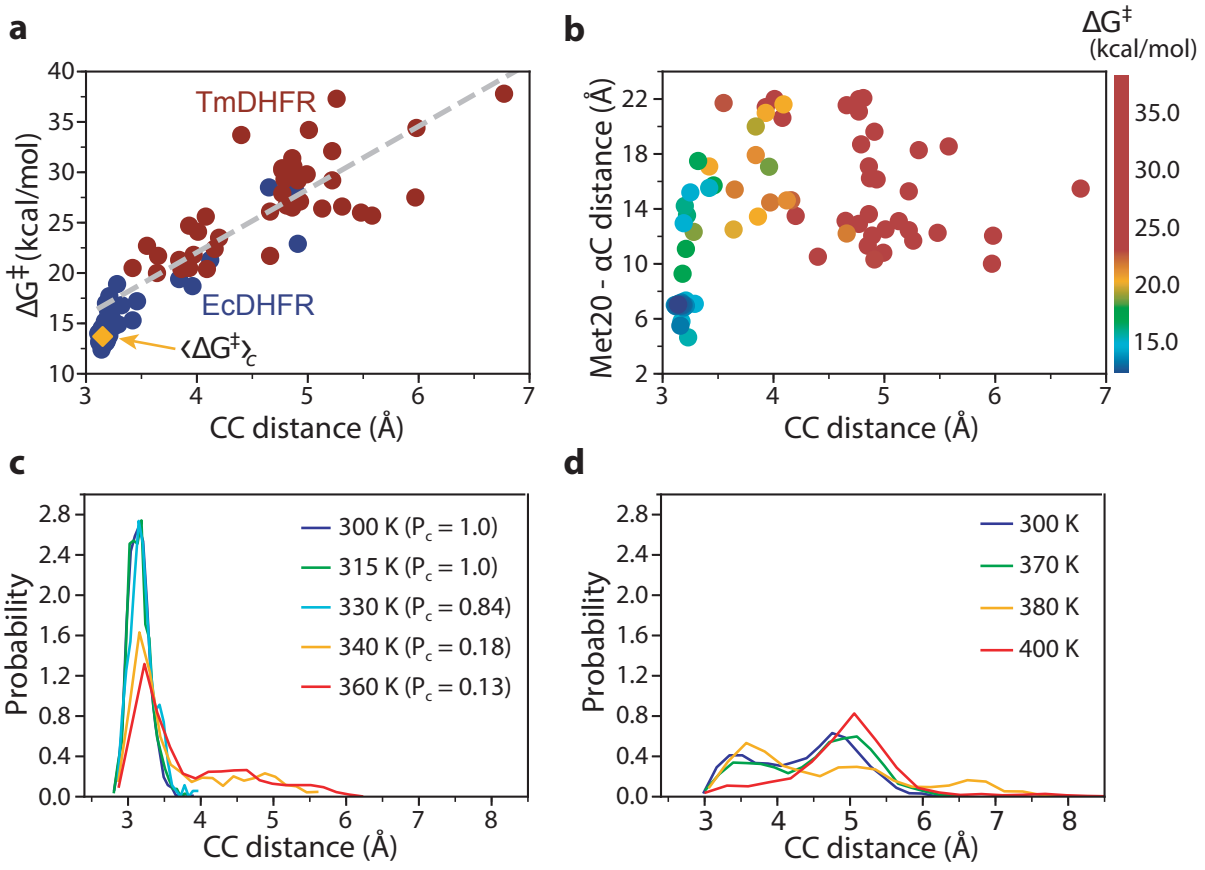

Figure 3. Effects of conformational changes on reactivity (**a**) Calculated $\Delta G^{\ddagger}$ as a function of CC distance in the reactant state for EcDHFR (blue) and TmDHFR (red). Each point corresponds to an EVB free energy profile simulated at a temperature within the investigated range (see Supplementary Information); the yellow diamond shows the average free energy barrier in the closed loop conformation $\langle \Delta G^{\ddagger} \rangle_c$. The gray dashed line corresponds to the linear regression used to estimate $\Delta G^{\ddagger}_{open}(CC)$ in equation (1). (**b**) Correlation between the Met20 loop conformation and the CC distance, where each point is colored according to the $\Delta G^{\ddagger}$ value obtained from a distinct EVB free energy calculation. (**c** and **d**) CC distributions at several temperatures below melting for EcDHFR (**c**) and TmDHFR (**d**). The fraction of closed loop conformations at each temperature is specified for EcDHFR and is always 0 for TmDHFR.

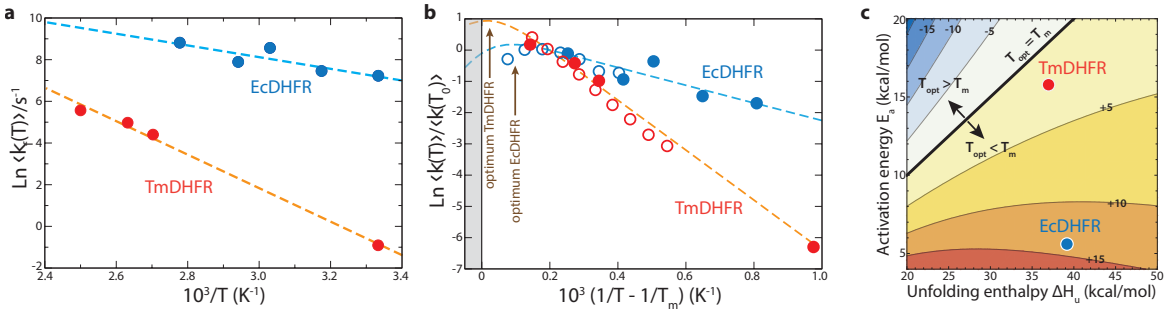

Figure 4. Comparison with experimental data and model for temperature-dependent activity (**a**) Arrhenius plot of $\langle k_f(T) \rangle$ for EcDHFR (blue circles) and TmDHFR (red circles), calculated using equation (1). The blue (Ec) and orange (Tm) dashed lines indicate the Arrhenius fits used to obtain E_a. (**b**) Arrhenius plots of the hydride transfer rate constant $\langle k(T) \rangle$ for EcDHFR and TmDHFR, respectively from experiments[30] (blue (Ec) and red (Tm) open circles), from our simulations (blue and red full circles, respectively), and from our model (dashed blue line and orange line, respectively) equation(3). The rate constant values are referenced to temperature T_0 such that $10^3(1/T_0\text{-}1/T_m) = 0.2$. The vertical dashed brown line indicates the position of the predicted optimal temperature T_{opt}, while the vertical black line corresponds to the melting temperature T_m. (**c**) 2D contour plot of $T_m - T_{\text{opt}}$ obtained from equation (5) using a fixed $T_m = 396$ K, as a function of activation energy and unfolding enthalpy (using a different T_m value leads to minor changes in $T_m - T_{\text{opt}}$ in equation (5)). The values corresponding to Ec and TmDHFR simulations are indicated by blue and red dots, respectively.

Supplementary Information

TABLE OF CONTENTS

I. Simulation methodology

1. Structure preparation

The X-ray structures of *E. Coli* DHFR (EcDHFR) in complex with folate and NADP+, of *T. Maritima* DHFR (TmDHFR) in complex with methotrexate and NADPH were used as initial structures (PDB codes 1RX2[1] and 1D1G[2], respectively). In order to generate the apo states, all the ligands eventually present were manually removed. Conversely, for the Michaelis-Menten complex (MM complex) of EcDHFR the 7,8-dihydrofolate (FOL) and the $NADP^+$ were replaced by the N5 protonated 7,8-dihydrofolate (FOL^+) and NADPH, respectively. For the MM complex of TmDHFR, the methotrexate was mutated to FOL^+ and its pterin ring moiety was flipped of 180°, in order to correctly reproduce the FOL binding pose.[3] The protonation state of the residues of the obtained systems was set at physiological conditions (pH = 7, salinity= 0.10 M) and hydrogens were added with the Chimera software[4]. The parameters for NADPH and FOL^+ were taken from the literature[5], protein atoms were described by the ff99SB Amber force field[6], and water molecules with the TIP3P[7] model.

Each system was solvated with a cubic box of water molecules ensuring that all protein atoms were at least 10 Å from the box edges, and the negatively charged proteins were neutralized by adding an adequate number of Na^+ ions.

2. Simulation parameters

Most simulations were performed with the NAMD 2.9 software[8]. The PME algorithm (grid spacing = 1 Å) was used to handle long-range contributions of electrostatic interactions, while a cutoff of 9 Å was set for short-range interactions and

real space contribution of electrostatic interactions. All bonds involving hydrogens were constrained.

3. Initial equilibration

After minimization, the systems were equilibrated under ambient conditions for 200 ns in the NPT ensemble using a Langevin thermostat (characteristic time 1 ps, T=300 K) and barostat (dumping time 50 fs, P = 1 atm) and an integration time of 2 fs.

4. REST2 simulations for conformational sampling

REST2 simulations were performed by using an in-house implementation in NAMD 2.9[9]. Within the REST2 scheme the replica evolve at a reference temperature β_{ref}, while the potential energy of the nth replica (E_n) is rescaled as:

$$E_n(\vec{X}) = \lambda_n E_{pp}(\vec{X}) + \sqrt{\lambda_n} E_{pw}(\vec{X}) + E_{ww}(\vec{X}) \quad \text{(S1)}$$

with $E_{pp}(\vec{X})$, $E_{pw}(\vec{X})$ and $E_{ww}(\vec{X})$ being the protein-protein, protein-solvent and solvent-solvent potential energies. Therefore, in each nth replica, the solvent evolves at the reference temperature, protein-solvent interactions at a temperature $\beta_n = \sqrt{\lambda_n}\beta_{ref}$, and protein-protein interactions at $\beta_n = \lambda_n \beta_{ref}$. For these latter, only the dihedral and the non-bonded degrees of freedom were rescaled, while protein bonds, angles and impropers are left unperturbed. This was ensured by rescaling the protein dihedral force constants and Lennard-Jones energies by λ_n and protein atomic charges by $\sqrt{\lambda_n}$ [10,11]. Because protein-protein and protein-solvent interactions are scaled differently, we have shown that an effective temperature $\langle \beta'_n \rangle$ can be defined for each replica using the corresponding state principle and a mean field approximation[11]:

$$\langle \beta_n' \rangle = \beta_n \left(1 + \left(\sqrt{\frac{\beta_{ref}}{\beta_n}} - 1 \right) \langle \frac{E_{pw}(\vec{X})}{E_{pw}(\vec{X}) + E_{pp}(\vec{X})} \rangle \right) \quad \text{(S2)}$$

For TmDHFR, only one of the two monomers was subjected to the rescaling scheme and the other one treated as solvent molecules. In the case of the MM complexes, the ligands were also treated as solvent to avoid unrealistic molecular geometries arising from potential energy rescaling. 24 replica exchanging protein-protein corresponding temperatures of 289, 300, 311, 323, 335, 347, 360, 373, 387, 402, 417, 432, 448, 465, 482, 500, 519, 538, 558, 579, 600, 625, 634, 652 K were used and the replica were allowed to exchange every 10 ps (success rate ~ 25%). The simulation protocol was similar to that used for the equilibration of the reactant state, except that atomic coordinates were output every 50 ps. Each simulation was run for 500 ns/replica, for a total of 12 μs. Overall, the replicas scanned an effective temperature window of $T_{eff} \in$ [292 K, 500 K]. Only the last 250 ns of the simulations on EcDHFR and TmDHFR were considered for the analysis. In the case of TmDHFR the analyses were only performed on the rescaled monomer.

In order to prevent the ligands from leaving the binding site, during the REST2 simulations of the MM complexes, harmonic restraints on three protein-FOL^+ and three protein-NADPH distances were included. These distances were selected by evaluating the most stable hydrogen bonds and hydrophobic interactions between the protein and the ligands along the 200-ns classical equilibration of each complex (see **Table S1**). The distances between the donor atom and the acceptor atom for each hydrogen bond, or between two atoms involved in the hydrophobic interaction were computed and used to define the equilibrium distance for the harmonic restraint as the

most occurring one in the relative distribution (**Table S1**). The harmonic force constant was set to 5 kcal·mol^{-1}·Å^{-2}.

Table S1 Atoms selected for REST2 distance harmonic restraints and their equilibrium distances.

EcDHFR			TmDHFR		
Atom 1	Atom 2	Eq. distance (Å)	Atom 1	Atom 2	Eq. distance (Å)
NPH - N7N	Ala7 - O	3.0	NPH – O'A5	Ile46 - N	2.7
NPH - O'N5	Gly97 - N	3.0	NPH – O'N5	Lys103 - N	3.0
NPH - OPA2	Thr46 - OG1	2.7	NPH - OPA2	Thr47- OG1	3.4
FOL$^+$ - O2	Arg57 - NH2	2.8	FOL$^+$ - C16	Phe31 - CZ	4.0
FOL$^+$ - N2	Asp27 - OD1	2.8	FOL$^+$ - O1	Arg58 – NH1	2.7
FOL$^+$ - N8	Ile5- O	3.0	FOL$^+$ - N8	Val6- O	3.1

5. EVB simulations for estimation of reaction free-energy barriers

We chose temperatures within the effective temperature range below the computed melting temperatures, namely 300.0, 314.8, 329.8, 337.6 and 362.3 K (300, 315, 330, 340 and 360 K in the main text) for EcDHFR and 300, 373.9, 383.7 and 403.1 K (300, 370, 380 and 400 K in the main text) for TmDHFR. From the corresponding trajectories of the MM complexes REST2 simulations we randomly selected 10 conformations for each temperature. Each conformation was equilibrated for 50 ns without distance restraints, using the protocol above described for the system equilibration, and at a physical temperature equivalent to the corresponding REST2 effective temperature. Each of the equilibrated conformation was then used as starting point for the calculation of the reaction free-energy. This was done by using the

Empirical Valence Bond (EVB) method[12–14] with the Amber12 package[15], as now described.

We used two valence bond diabatic states, hereby termed VB1 and VB2. The former corresponds to the reactant state, composed of the protonated folate and the cofactor, while the latter reproduces the product state, consisting of the tetrahydrofolate and the oxidized $NADP^+$. For such a system, the Hamiltonian can be written as:

$$\widehat{H} = \begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} + \Delta \end{bmatrix} \quad \text{(S3)}$$

where V_{ii} is the potential of VB state i, Δ is the constant offset between the two states, and V_{12} is the electronic coupling between the two states. The adiabatic ground state energy of the system corresponds to the lowest eigenvalue of this Hamiltonian. We used the constant offset (-60.86 kcal/mol) and the electronic coupling (44.15 kcal/mol) fit by Hammes-Schiffer and coworkers[5] to reproduce the experimental free energy of activation and free energy of reaction of the EcDHFR.

For each VB state the whole system was described classically by the ff99SB Amber force field[6], except for the C-H harmonic bond which is broken and created. This was described by a Morse potential written as:

$$V_{Morse}(r_{ij}) = D_e \left[1 - e^{-\alpha\left(r_{ij} - r_{ij}^0\right)}\right]^2 \quad \text{(S4)}$$

where D_e is the potential depth, r_{ij} is the distance between the C donor or C acceptor and the transferred H, r_{ij}^0 is this distance at the equilibrium and $\alpha = \sqrt{k_e/2D_e}$, with k_e being the force constant at the potential minimum. In our

simulations we used the parameters reported by before[5], with D_e set to 103.0 kcal/mol, r_{ij}^0 to 1.09 Å and α to 1.817 Å^{-1}.

The EVB calculations were carried out with Amber12 in the NVT ensemble, using a Langevin thermostat with a collision frequency of 0.1 ps^{-1}, the SHAKE algorithm to constrain the bonds involving hydrogens, except that involving the transferred hydride, and a nonbonded cutoff of 13 Å and PME algorithm for long-range interactions. Van der Waals interaction between donor-hydride and acceptor-hydride were removed.

The calculation of the barrier free energy for the hydride transfer reaction ($\Delta G^{\ddagger}$) required the computation of the free energy profile along the reaction coordinate (**Figure S1B**). This was set to be the instantaneous energy gap between the two diabatic states ($\Delta E = V_{11} - V_{22}$) (Figure S1**A**), which is negative in the reactant state, zero at the transition state and becomes positive moving toward the product state.

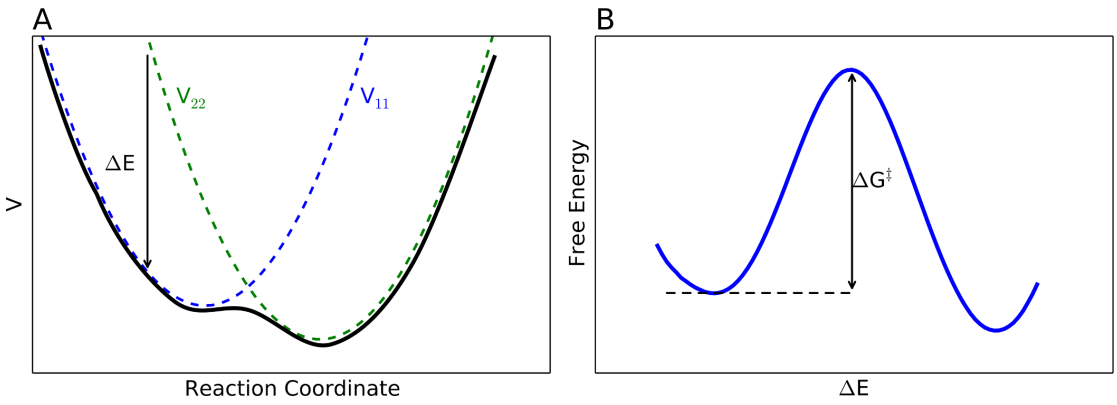

Figure S1. (A) EVB potential energy as a function of the environmental reaction coordinate, with the groundstate potential indicated as full black line. The potential of the first VB state (V_{11}) and of the second VB state (V_{22}) are represented as dashed blue and green lines, respectively. The reaction coordinate chosen for the

generation of the hydride transfer free energy profile is the instantaneous energy gap (ΔE) between V_{22} and V_{11}. (B) Schematic representation of the reaction free energy profile as a function of the reaction coordinate ΔE. The barrier free energy is labeled as $\Delta G^{\ddagger}$.

Since the experimental reaction free energy for the EcDHFR catalyzed hydride transfer is much higher than thermal fluctuations (around 13 kcal/mol and 0.6 kcal/mol, respectively), an enhanced-sampling scheme was employed based on a mapping potential approach, which constraints the system to a potential V_λ being a fixed mixture of the two VB potentials:

$$V_\lambda = (1 - \lambda)V_{11} + \lambda V_{22} \quad \text{(S5)}$$

The modification of λ allows bringing the system from the reactant state ($\lambda \sim 0$) to the product state ($\lambda \sim 1$), via the transition state ($\lambda \sim 0.5$).

Only half of the free energy profile is needed to compute the barrier free energy. Therefore, from each equilibrated conformation half of the free energy profile was generated using 18 windows with a $\Delta\lambda = 0.025$ and a starting λ of 0.075. At each increment, the system was equilibrated for 15 ps before moving to the next window. Then, in each window a 100 ps production run was performed. For each of the obtained trajectories, half of the free energy profile along the reaction coordinate was reconstructed with the EVB groundstate without constraint using the Weighted Histogram Analysis Method (WHAM). From the obtained half-profiles, the activation free energy was calculated.

II. Necessity of an enhanced sampling approach for conformational space exploration

Brute force MD simulations were performed at the experimental melting temperatures evidence these challenges. For each system, we started from the crystal structure propagated for microsecond at the experimental melting temperatures (326 K[16] for EcDHFR and 356 K[17] for TmDHFR). As expected, no noticeable change of the proteins structure was observed on this timescale (**Figure S2** and **Figure S3**), with atomistic fluctuations of the protein backbone very close to those observed at 300~K (**Table S2**). By contrast, experimental data shows that at this temperature, half of the enzymes should populate non-native conformations, which is clearly not the case in our simulations. Our own enhanced sampling simulations, that we reported recently, suggest that the enzyme structures around melting are significantly different from the native state (**Table S2**).

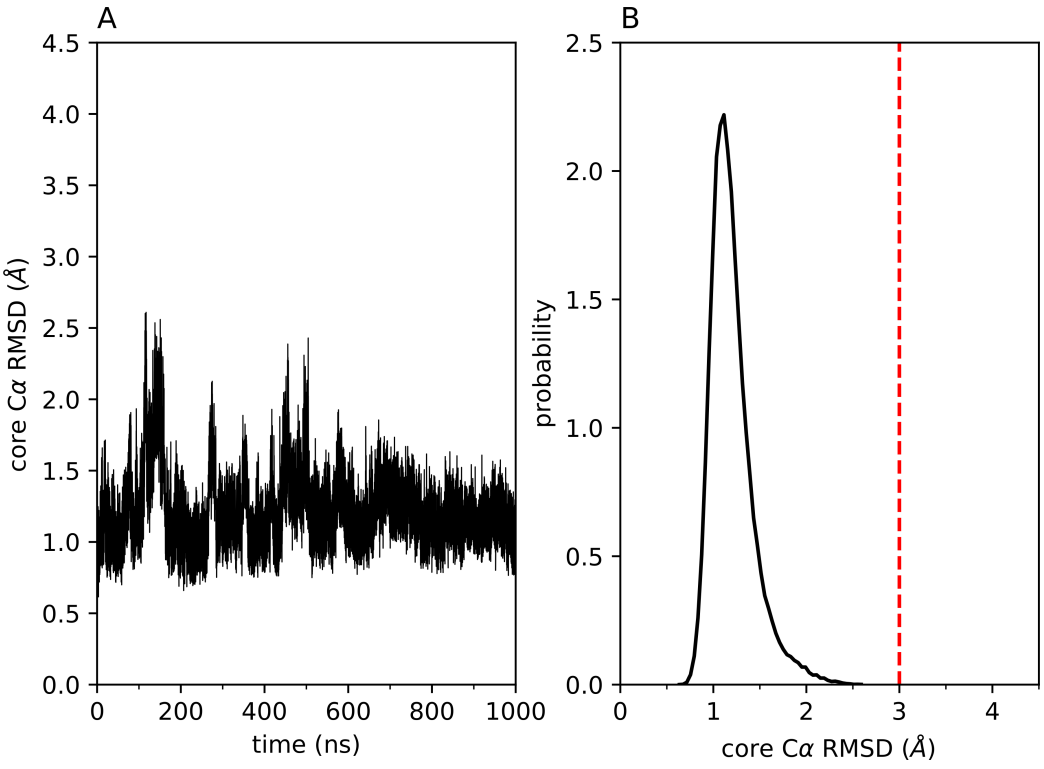

Figure S2. Root mean-square displacement (RMSD) computed on all non-loops Cα with respect to the crystal structure as a function of simulation time for EcDHFR at its experimental melting temperature T=323 K as a function of time (A), and distributions averaged over the 1 μs-long simulation (B). Averages reported in **Table S2**.

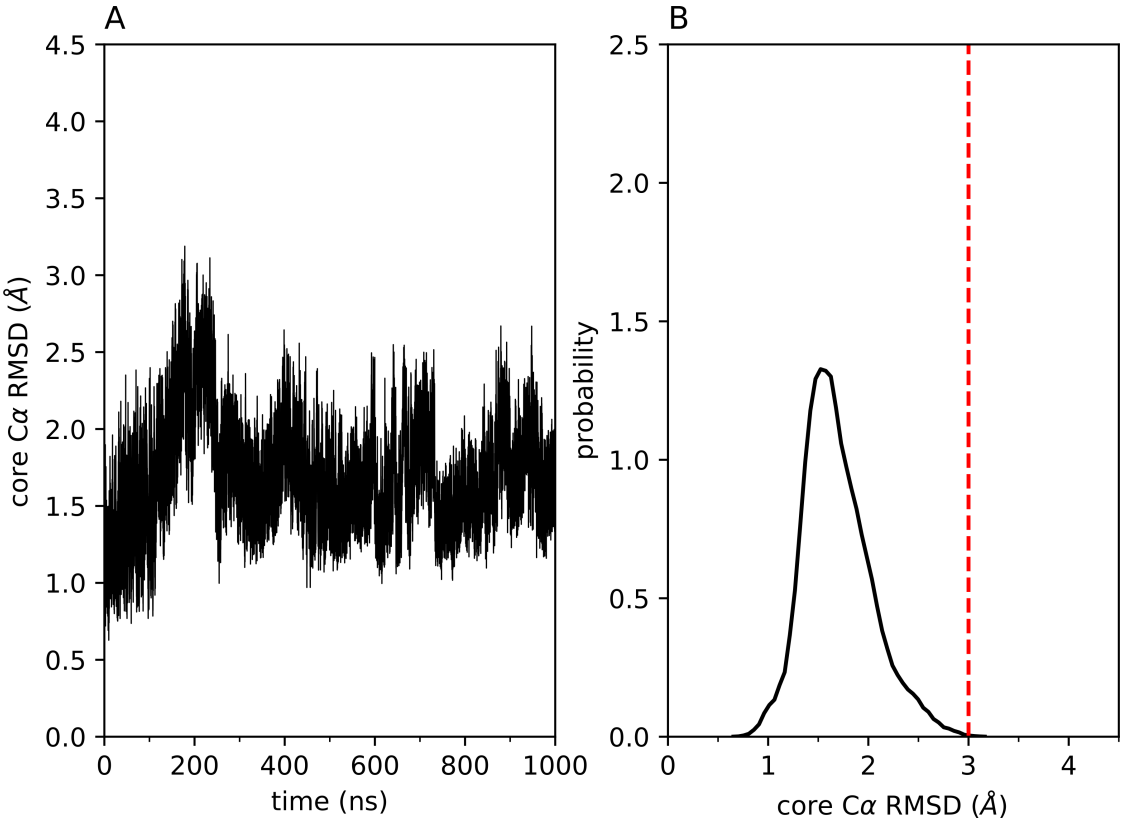

Figure S3. RMSD as a function of simulation time for TmDHFR at its experimental melting temperature T=357 K as a function of time (A), and distributions averaged over the 1 μs -long simulation (B). Averages reported in **Table S2**.

Table S2. Average RMSD and standard deviations for the microsecond-long simulations performed at the experimental melting temperatures (**Figure S2** and **Figure S3**), the REST2 replica at an effective temperature of 300 K, and the REST2

replicas that are the closest to the melting temperature determined from the simulated stability curves[18].

	EcDHFR	TmDHFR
Brute force MD, exp Tm	1.20 ± 0.24 Å	1.69 ± 0.34 Å
REST2 – 300K replica	1.14 ± 0.29 Å	0.97 ± 0.13 Å
REST2 - calc Tm replica	3.51 ± 0.79 Å	3.87 ± 0.62 Å

III. Relationship between Met20 loop conformation, CC distance and free-energy for hydride transfer

1. Met20 loop conformations and CC distributions

We discriminated between Met20 loop open and close conformations by analyzing the distance between this loop and the αC helix (loop-helix distance). In particular, we looked at the distance between the EcDHFR Asn18 and His45 Cα atoms, which, after structural alignment, corresponds to the TmDHFR Val19 and Ile46 Cα atoms. If this distance is between 6 and 8 Å the Met20 loop is closed, while if it is longer than 15 Å the loop is open. Indeed, in the crystallographic structure of EcDHFR, where the loop is closed, the distance between Asn18 and His45 Cα atoms is 6.9 Å, while in the X-ray structure of TmDHFR, where the loop is open, the distance between Val19 and Ile46 Cα atoms is 18.4 Å. In addition, it has been experimentally observed that in the EcDHFR closed Met20 loop conformation Asn18 and His45 are in close contact with the possibility of having a H-bond between the Asn18 side chain nitrogen and the His45 backbone oxygen[1]. Conversely, in the Met20 loop open conformation it has been observed that the Val19 side chain creates

hydrophobic interactions with Tyr125,[2] which belongs to the loop between the βF and βG strands (Figure 1 in the main text) and is more than 14 Å from His45 and Ile46 in EcDHFR and TmDHFR respectively.

When temperature increases, the Met20 loop of EcDHFR starts to sample more open conformations, which affects the CC distance in the reactant state (RS) which becomes longer (**Figure S4** and main text Figure 3). This results in a distortion of the CC distance in the transition state (TS) configurations as well, as shown in **Figure S4**, which then leads to higher reaction free-energy barriers, as shown further.

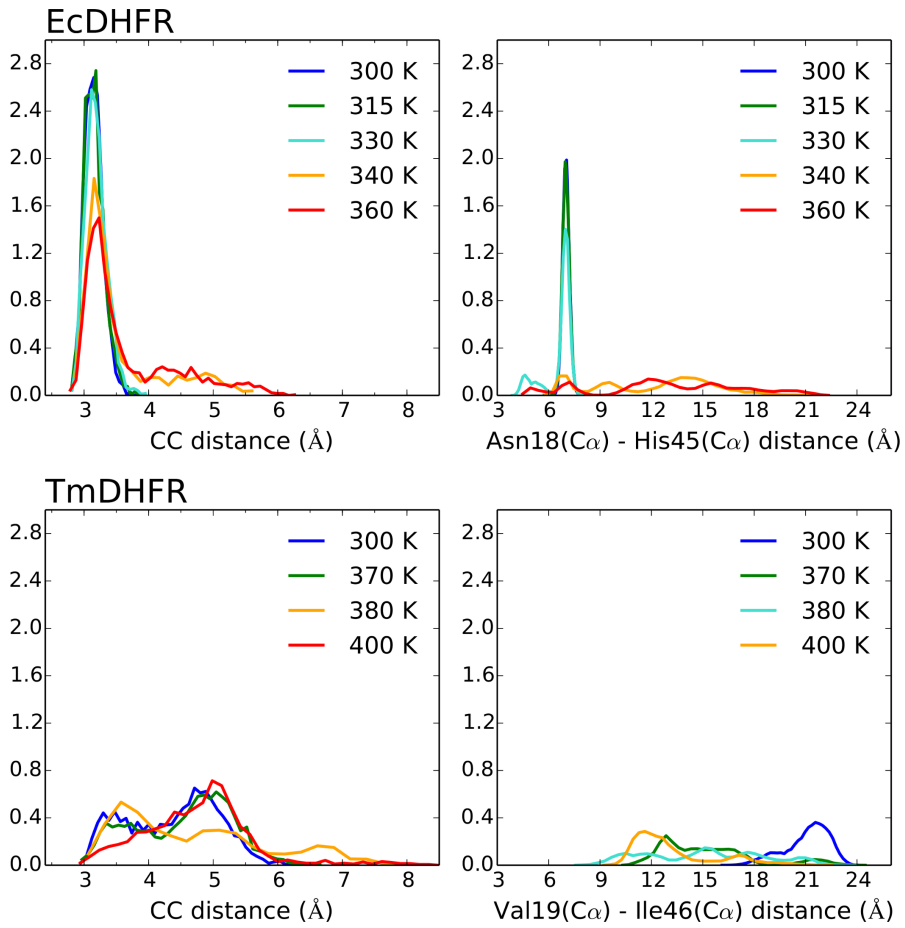

Figure S4. Reactant state (λ= 0.075, 0.10) CC distance distributions (left panels) and Met20 loop - αC helix distance distributions (right panels) for EcDHFR

(top panels) and TmDHFR (bottom panels) at different temperatures from EVB calculations. For EcDHFR at low temperatures (below 330 K) the closed conformation, indicated by a Asn18(Cα)-His45(Cα) distance between 6 and 8 Å, allows sampling short CC distances. Temperature increase brings the system to sample Met20 loop conformations other then closed and short CC distances are not ensured anymore. In the case of TmDHFR, the closed Met20 loop conformation is never sampled, as showed by Val19(Cα)-Ile46(Cα) distances greater than 8 Å. Globally longer CC distances are therefore sampled.

Figure S5 reports the transition state CC distance distributions in EcDHFR and TmDHFR. These distributions do not exhibit significant changes with temperature and with the homolog, in agreement with recent calculations[19]. The average transition state CC distances found here are in excellent agreement with prior calculations[20] for EcDHFR at 300K performed with the same valence bond state model and with the same collective energy gap coordinate, but slightly longer than reported in another set of calculations using a QM/MM approach and the hydride position as a reaction coordinate[19].

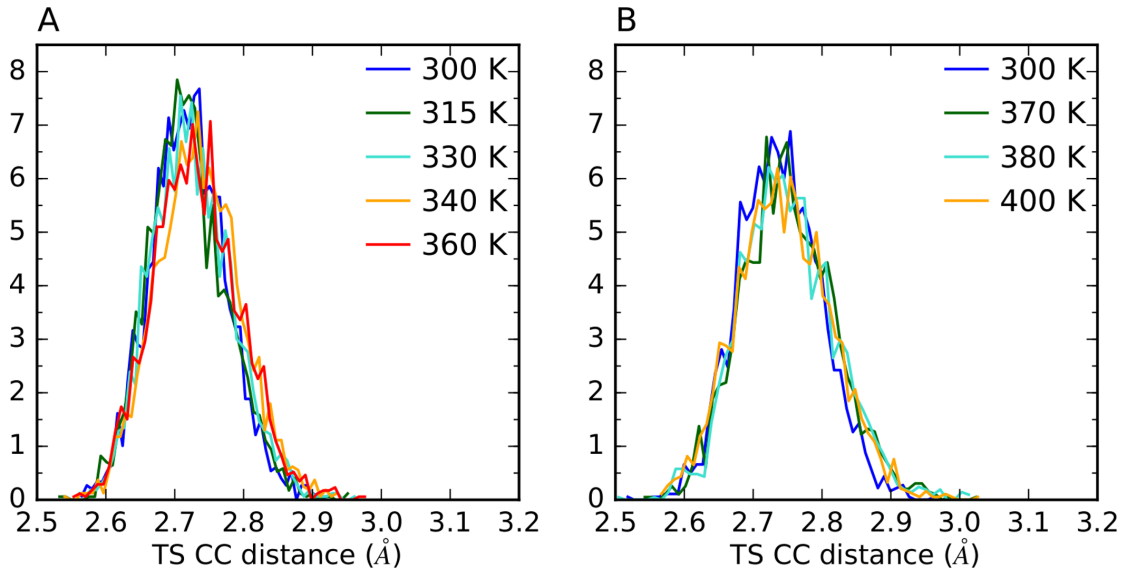

Figure S5. Transition state (λ= 0.5) CC distance distributions for EcDHFR (A) and TmDHFR (B) at different temperatures from EVB calculations.

2. Stabilizing interactions in the active site upon Met20 loop closure

The Met20 loop conformation has key consequences for the interactions between the active site residues and the substrates, as shown in **Figure S7** for a set of relevant distances defined in **Figure S6**.

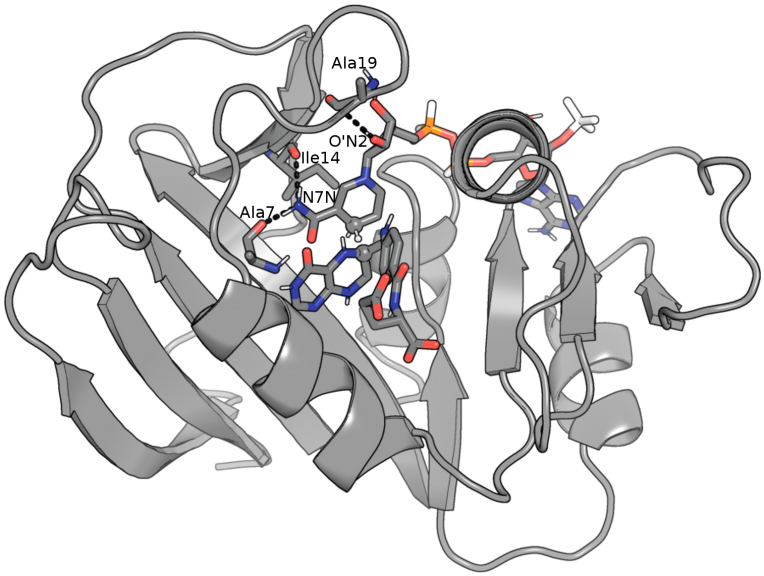

Figure S6. X-ray structure of EcDHFR (PDB code 1R2X) in complex with NADPH and folate (represented as sticks). The carbons donating and accepting the hydride and the hydride are reported as spheres. The hydrogen bonds between NADPH and the protein observed in the closed conformation are represented as black dashed lines, and the involved NADPH atoms and protein residues are labeled.

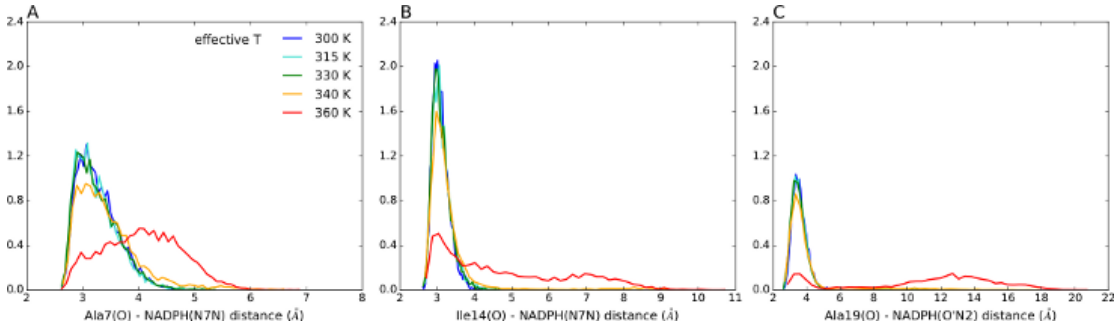

Figure S7. Atom donor – atom acceptor distance distributions at 300 (blue), 315 (turquoise), 330 (green), 340 (yellow) and 360 K (red) for relevant hydrogen bonds between the NADPH cofactor and the Met20 loop. If the Met20 loop is in a closed conformation the three reported H-bonds (Ala7(O)-NADPH(N7N), Ile14(O)-

NADPH(N7N) and Ala19(O)-NADPH(O'N2)) help in stabilizing the NADPH orientation respect to the folate substrate.

3. *Effect of Met20 loop conformation and CC distance on* $\Delta G^{\ddagger}$

The reaction free energy barrier $\Delta G^{\ddagger}$ is calculated in EcDHFR and TmDHFR at a series of temperatures, in 10 independent configurations at each temperature, leading to a total of 90 configurations (**Table S3** and **Table S4**). The $\Delta G^{\ddagger}$ values are well described by a bilinear fit as a function of the reactant CC distance and of the temperature,

$$\Delta G^{\neq} = \alpha + \beta\, CC + \gamma\, T \qquad \text{(S10)}$$

with α= -9.4454 kcal/mol, β=7.14378 kcal/(mol.Å), and γ = 0.005639 kcal/(mol.K). The resulting root mean square error is 2.77 kcal/mol. The effect of the CC distance on $\Delta G^{\ddagger}$ is further illustrated in **Figure S8**.

The α term can be interpreted as an activation enthalpy $\Delta H^{\ddagger}$, which is 12.7 kcal/mol at CC=3.1 Å, and γ as $-\Delta S^{\ddagger}$, which leads to a $-T\,\Delta S^{\ddagger}$ activation entropy contribution to the free energy barrier of approximately 1.7 kcal/mol at 300 K.

The amplitude of the change in $-T\,\Delta S^{\ddagger}$ over the investigated 300 - 400K temperature range is therefore 0.6 kcal/mol, which is negligible with respect to the 28.6 kcal/mol amplitude of the change in activation enthalpy induced by the 3 - 7 Å CC fluctuations.

This shows that changes in $\Delta G^{\ddagger}$ are mostly due to changes in the CC distance.

Table S3. Average Asn18(Cα)-His45(Cα) and CC distance and $\Delta G^{\ddagger}$ for each EVB run on EcDHFR at 300, 315, 300 340 and 360 K. The Asn18(Cα)-His45(Cα)

allows to define the Met20 loop conformation, which is closed when this distance is between 6 and 8 Å. The CC distance refers to the CC distance averaged over the windows corresponding to the reactant states (λ = 0.075 and 0.1).

EcDHFR, 300 K			
run	Asn18(Cα)-His45(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	7.04 ± 0.19	3.15 ± 0.13	13.5
2	7.03 ± 0.18	3.19 ± 0.15	14.0
3	7.00 ± 0.20	3.11 ± 0.13	14.0
4	6.96 ± 0.20	3.18 ± 0.15	13.2
5	7.04 ± 0.20	3.14 ± 0.14	12.4
6	7.00 ± 0.20	3.15 ± 0.14	13.8
7	7.09 ± 0.20	3.18 ± 0.14	13.9
8	6.95 ± 0.20	3.16 ± 0.14	13.3
9	6.94 ± 0.22	3.12 ± 0.14	13.2
10	7.08 ± 0.20	3.14 ± 0.14	13.5
EcDHFR, 315 K			
run	Asn18(Cα)-His45(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	6.98 ± 0.22	3.19 ± 0.17	14.4
2	7.00 ± 0.20	3.17 ± 0.17	14.2
3	7.07 ± 0.21	3.14 ± 0.14	14.4
4	7.06 ± 0.20	3.17 ± 0.15	13.7
5	6.91 ± 0.23	3.12 ± 0.14	14.1
6	6.99 ± 0.21	3.15 ± 0.14	14.6
7	7.02 ± 0.20	3.18 ± 0.16	13.1
8	7.07 ± 0.21	3.17 ± 0.16	13.3
9	7.01 ± 0.21	3.14 ± 0.15	13.9
10	6.98 ± 0.22	3.19 ± 0.17	14.4
EcDHFR, 330 K			
run	Asn18(Cα)-His45(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	7.00 ± 0.20	3.15 ± 0.16	14.1
2	5.50 ± 0.47	3.16 ± 0.15	13.5
3	6.88 ± 0.24	3.18 ± 0.15	14.0
4	7.01 ± 0.23	3.14 ± 0.17	14.5
5	7.02 ± 0.22	3.19 ± 0.17	13.3
6	6.95 ± 0.21	3.21 ± 0.17	13.8

7	7.07 ± 0.21	3.18 ± 0.17	14.1
8	7.09 ± 0.21	3.29 ± 0.17	14.9
9	4.65 ± 0.26	3.23 ± 0.18	15.6
10	6.99 ± 0.28	3.16 ± 0.14	12.9
EcDHFR, 340 K			
run	Asn18(Cα)-His45(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	9.28 ± 0.56	3.18 ± 0.15	17.0
2	17.07 ± 1.7	3.96 ± 0.39	18.7
3	6.82 ± 0.26	3.19 ± 0.15	13.7
4	14.22 ± 1.03	3.20 ± 0.17	16.2
5	15.2 ± 1.11	3.25 ± 0.19	14.7
6	10.32 ± 1.16	4.91 ± 0.32	22.9
7	13.54 ± 0.77	3.22 ± 0.18	16.4
8	7.33 ± 0.86	3.21 ± 0.16	15.2
9	12.96 ± 1.58	3.19 ±0.15	15.1
10	14.62 ± 1.14	4.12 ± 0.68	21.3
EcDHFR, 360 K			
run	Asn18(Cα)-His45(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	12.06± 1.35	4.89 ± 0.73	28.1
2	15.71 ± 1.30	3.46 ± 0.42	17.2
3	11.09 ± 0.73	3.21 ± 0.16	17.7
4	20.00 ± 0.99	3.84 ± 0.64	19.4
5	5.81 ± 1.00	3.17 ± 0.16	15.2
6	7.19 ± 0.38	3.17 ± 0.17	13.6
7	15.53 ± 0.55	3.79 ± 0.58	15.3
8	13.13 ± 0.75	4.65 ± 0.43	28.5
9	17.49 ± 1.05	3.32 ± 0.20	16.8
10	12.33 ± 0.55	3.28 ± 0.17	18.9

Table S4. Average Val19(Cα)-Ile46(Cα) and CC distance and $\Delta G^{\ddagger}$ for each EVB run on TmDHFR at 300, 370, 380 and 400 K. The Val19(Cα)-Ile46(Cα) allows to define the Met20 loop conformation, which is closed when this distance is between

6 and 8 Å. The CC distance refers to the CC distance averaged over the windows corresponding to the reactant states (λ = 0.075 and 0.1).

TmDHFR, 300 K			
run	Val19(Cα)-Ile46(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	21.09 ± 0.65	4.77 ± 0.60	27.9 (shoulder)
2	21.71 ± 0.69	3.55 ± 0.35	22.7
3	22.00 ± 0.70	4.02 ± 0.44	24.1
4	20.63 ± 0.83	4.08 ± 0.55	25.6
5	18.70 ± 0.74	4.79 ± 0.33	29.3
6	21.41 ± 0.80	3.93 ± 0.59	24.7
7	21.55 ± 0.87	4.66 ± 0.95	26.1
8	19.62 ± 0.93	4.91 ± 0.33	29.2
9	21.99 ± 0.62	4.77 ± 0.35	30.2
10	22.07 ± 0.80	4.81 ± 0.36	26.7
TmDHFR, 370 K			
run	Val19(Cα)-Ile46(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	15.28 ± 0.73	5.22 ± 0.30	29.2
2	12.45 ± 0.52	5.22 ± 0.54	32.1
3	12.91 ± 0.37	4.77 ± 0.28	30.4
4	16.15± 1.27	4.93 ± 0.52	27.1
5	12.51 ± 0.95	5.01 ± 0.24	34.2
6	14.47 ± 0.60	3.97 ± 0.36	21.8
7	21.61 ± 0.84	4.09 ±0.71	20.4
8	16.23 ± 0.70	4.87 ± 0.47	30.6
9	17.09 ± 0.65	3.42 ± 0.23	20.5
10	13.44 ± 0.53	3.86 ± 0.49	20.4
TmDHFR, 380 K			
run	Val19(Cα)-Ile46(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	18.28 ± 1.53	5.31 ± 0.67	26.6
2	12.50 ± 1.08	3.64 ± 0.34	20.0
3	15.41 ± 0.52	3.65 ± 0.35	21.7
4	15.48 ± 0.74	6.77 ± 0.40	37.8
5	17.92 ± 0.89	3.84 ± 0.40	21.3
6	10.01 ± 0.93	5.97 ± 0.78	27.5
7	20.99 ± 0.69	3.93 ± 0.47	20.5
8	14.63 ± 0.63	4.16 ± 0.67	22.4
9	10.52 ± 0.53	4.40 ± 0.58	33.7
10	18.6 ± 0.79	5.48 ± 1.07	26.0
TmDHFR, 400 K			
run	Val19(Cα)-Ile46(Cα) (Å)	CC distance (Å)	$\Delta G^{\ddagger}$ (kcal/mol)
1	12.05 ± 0.88	5.98 ±0.44	34.4
2	11.69 ± 0.60	5.26 ± 0.38	37.3

3	11.32 ± 0.46	4.85 ± 0.42	30.6
4	13.11 ± 0.89	5.13 ± 0.49	26.4
5	12.21 ± 0.72	4.66 ± 0.76	21.7
6	13.49 ± 1.00	4.20 ± 0.89	23.5
7	13.62 ± 1.75	4.86 ± 0.37	26.5
8	17.10 ± 0.57	4.86 ± 0.52	31.4
9	18.55 ± 1.37	5.58 ± 1.49	25.7
10	10.81 ± 0.46	4.99 ± 0.45	29.8

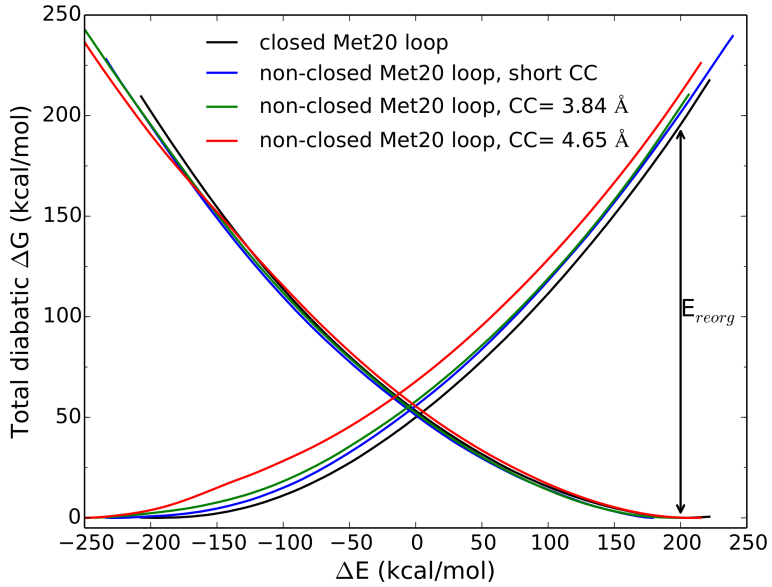

Figure S8. Diabatic free energy functionals at 360 K for EcDHFR in the closed conformation (black, run 6 **Table S33**), in non-closed conformation and short CC distance (blue, run 3 **Table S33**), in non-closed conformation and CC distance = 3.84 Å (green, run 4 **Table S33**) and in non-closed conformation and CC distance = 4.65 Å (red, run 8 **Table S33**). The reorganization energy (E_{reorg}) represents the reduction in the energy of the system when the reacting system is placed on the potential surface of the product state, at the equilibrium coordinate of the reactant state, and then let relax to the product equilibrium coordinate. The minimum E_{reorg} is obtained when the Met20 loop is closed, while, when it is in a non-closed conformation, it increases with the CC distance.

IV. Models to recover the effective reaction free-energy barriers from the simulations results

The EVB calculations revealed that the reaction free energy $\Delta G^{\ddagger}$ is very sensitive to the Met20 loop conformation through the CC distance. Provided with 10 simulations at each considered temperature (**Table S33** and **Table S44**), we therefore investigated several models in order to recover the average, effective barrier $\Delta G^{\ddagger}_{eff}$ that can be experimentally obtained with activity measurements (**Figure S9** and discussion below).

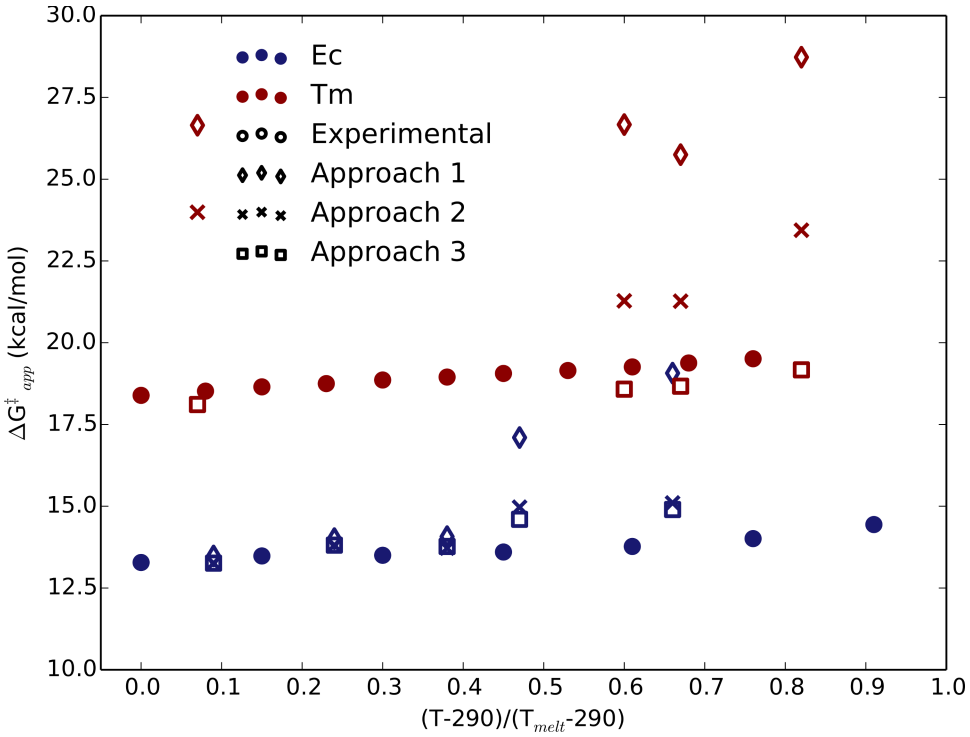

Figure S9. Comparison between experimental apparent barriers $\Delta G^{\ddagger}_{app}$ (filled circles), computed barriers $\Delta G^{\ddagger}$ averaged over 10 conformations (static average, open diamonds), barriers obtained by averaging over individual rates of each of the 10

selected conformations (dynamic average, crosses) and the results of our model (model, open squares) for EcDHFR (blue) and TmDHFR (red). Data are showed as a function of a rescaled temperature axis, which takes into account the differences between experimental and calculated melting temperatures.

1. *Barrier averaged over 10 conformations*

We first consider the ensemble average of the barriers over the 10 selected conformations, i.e. $\Delta G^{\ddagger}_{eff} = \langle \Delta G^{\ddagger} \rangle$. As shown in **Figure S9**, this approach largely overestimates the effective barrier, by overweighting conformations with large barriers.

2. *Chemical rate averaged over 10 conformations*

If we assume a fast equilibrium between all reactant conformations, the observed reaction rate is an average over the individual rates of each conformation, i.e., $k_{eff} = \langle k \rangle$. Indeed, assuming a conformational coordinate R fast as compared to the reaction coordinate, the reaction rate k is equal to $\frac{\int dR \times e^{-\frac{G^{\ddagger(R)}}{k_B T}}}{\int dR \times e^{-\frac{G^{RS(R)}}{k_B T}}}$, with $G^{RS(R)}$ and $G^{\ddagger(R)}$ being the free energies as a function of R in the reactant and transition state, respectively. This is equivalent to $k = \frac{\int dR \times e^{-\frac{\left(G^{RS(R)} + \Delta G^{\ddagger}(R)\right)}{k_B T}}}{\int dR \times e^{-\frac{G^{RS(R)}}{k_B T}}}$, which can be written as $\frac{\int dR \times p^{RS}(R) e^{-\frac{\Delta G^{\ddagger}(R)}{k_B T}}}{\int dR \times p^{RS}(R)}$, with $p^{RS}(R)$ being the probability of R in the reactant state. It follows that the observed effective barrier is equal to $-k_B T ln\left(\langle e^{-\frac{\Delta G^{\ddagger}_{C}}{k_B T}} \rangle\right)$. As shown in **Figure S9**, the effective barrier is still large as compared to the experimental value.

Indeed, the average is very sensitive to the population having low barriers, for which the rate is much faster, and it suffers from insufficient sampling of this region (which cannot be ensured by selecting 10 conformations to represent the entire CC distribution).

3. Chemical rate averaged over the CC distribution

To solve this issue, we propose a continuous model based on the CC distributions. We consider separately the population of closed conformations $P_c^{loop}(T)$ that always lead to a low, average barrier of $\langle \Delta G^{\ddagger} \rangle_c$ = 13.7 kcal/mol, and the temperature-dependent distribution of non-closed conformations $P_{nc}^{CC}(T)$ (Fig. 3C and D main text), where a linear relationship $\Delta G_{reg}^{\ddagger}$ between the barrier and the CC distance is observed. The resulting apparent barrier for the reaction $\Delta G_{app}^{\ddagger}$ (see above) is:

$$\Delta G_{app}^{\ddagger}(T) = -k_B T ln \left[P_c^{loop}(T) e^{-\langle \Delta G^{\ddagger} \rangle_c / k_B T} + \left(1 - P_c^{loop}(T)\right) \int dCC \times P_{nc}^{CC}(T) e^{-\langle \Delta G^{\ddagger} \rangle_{reg}(CC)/k_B T} \right] \quad \text{(S11)}$$

As shown in the main text and in **Figure S9**, this model is successful in reproducing the experimental values of the apparent free-energy barrier for EcDHFR and TmDHFR along a wide temperature range.

4. Assumption of fast equilibrium

We make the assumption that at a given temperature, conformational changes of the Met20 loop are faster than hydride transfer, which is legitimized by the fact that the barrier is large (~13 kcal/mol) compared to that of small conformational changes,

and by the fast decay of the CC time correlation function, as illustrated below (**Figure S10**).

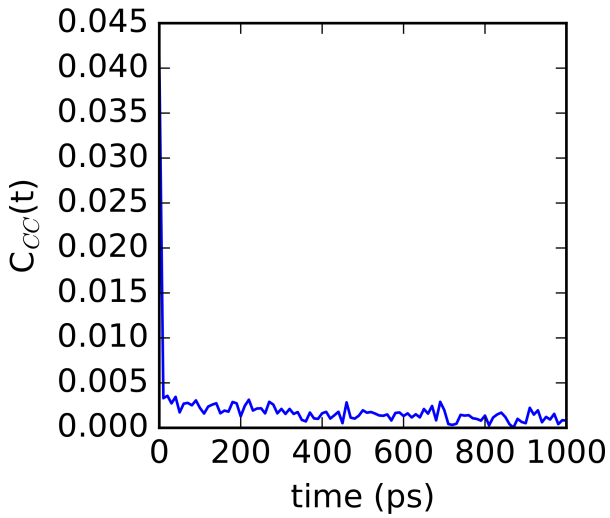

Figure S10. Time correlation function of CC distance in the reactant state ($C_{CC}(t)$).

V. Mechanisms of thermal stability and effect on optimal and melting temperatures

The Gibbs-Hemholtz equation provides the protein stability curve.

$$\Delta G_u = \Delta H_u(1 - T/T_m) - \Delta C_p[T(\ln T/T_m - 1) + T_m], \qquad \text{(S12)}$$

where ΔG_u is the unfolding free energy, ΔH_u is the unfolding enthalpy, ΔC_p is the unfolding heat capacity, and T_m is the melting temperature. The stability curve $\Delta G_u(T)$ is plotted in **Figure S11** for EcDHFR. It exhibits a maximum at the maximum stability temperature T_S, where the unfolding free energy is ΔG_s, and it vanishes at the melting temperature ΔG_u (T_m) = 0. The melting curve is locally parabolic around T_S.

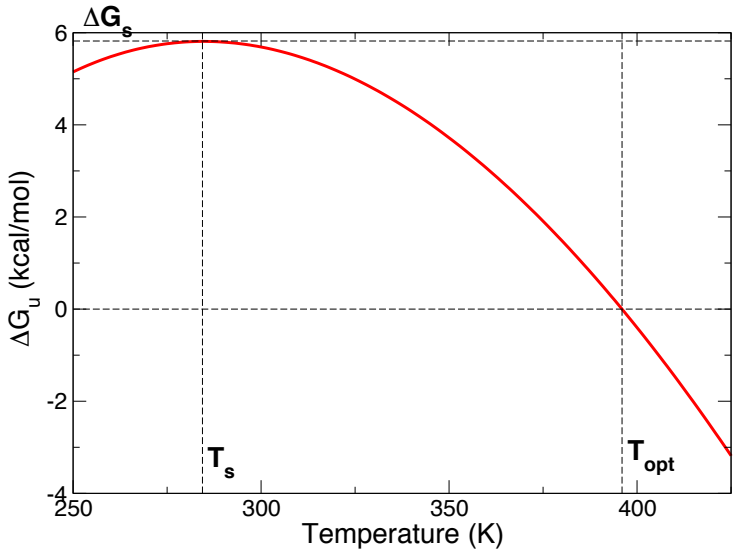

Figure S11. Stability curve of EcDHFR

Studies of protein thermal adaptation usually characterize stability curves via the location of the maximum stability (i.e. ΔG_s and T_s) and their width[21]. There are

thus three main different ways to increase the protein stability: upshifting the stability curve, i.e. increasing ΔG_s while keeping T_S and the width of the parabola constant, right shifting the stability curve, i.e. increasing T_s while keeping ΔG_s and the width fixed, broadening the stability curve, i.e. increasing the width, while keeping ΔG_s and T_s fixed.

We have thus considered these three possibilities and studied the impact on T_m and T_{opt} via our model, while keeping the reaction activation energy constant. In each case, we used EcDHFR as a starting point, determined (T_m, ΔH_u, ΔC_p), and then T_{opt} by solving Eq. S9 numerically. We explored ranges of values for ΔG_s, T_S and width such that the stability was increased to reach the melting temperature of TmDHFR. The results in **Figure S12** show that T_m and T_{opt} shift almost in parallel. This evidences that the significant difference in T_m-T_{opt} shifts between Ec and TmDHFR is mostly due to the different activation energies.

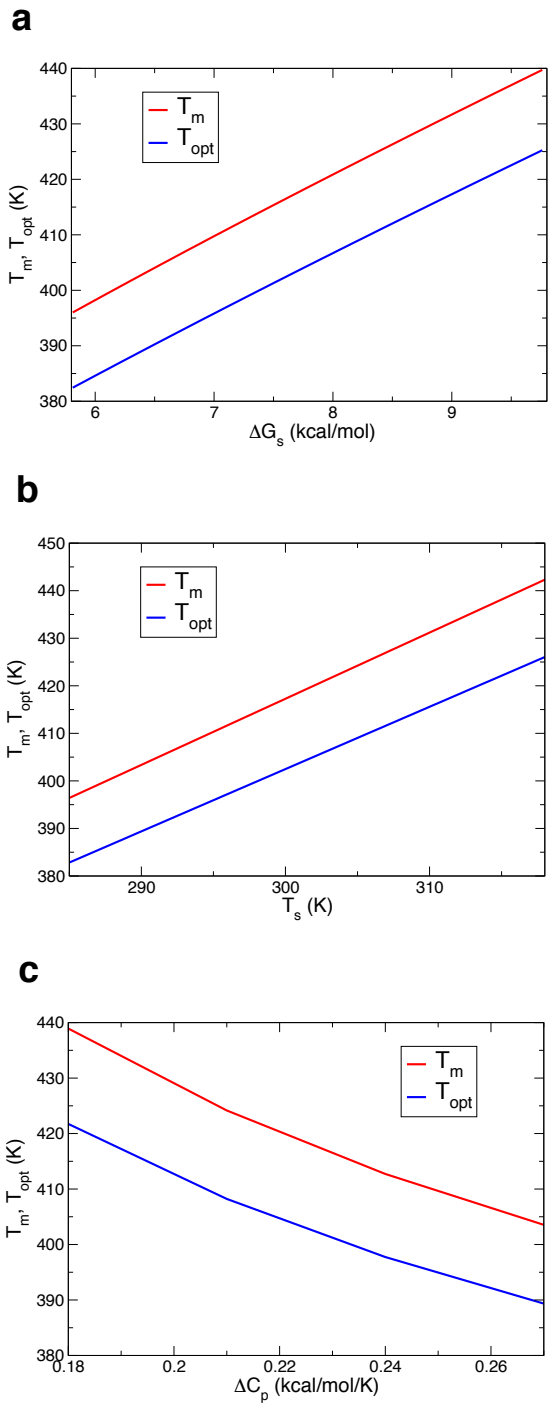

Figure S12. Melting and optimum activity temperatures for a stability curve a) upshift, b) rightshift and c) broadening

References

(1) Sawaya, M. R.; Kraut, J. Loop and Subdomain Movements in the Mechanism of Escherichia Coli Dihydrofolate Reductase: Crystallographic Evidence † , ‡. *Biochemistry* **1997**, *36*, 586–603.

(2) Dams, T.; Auerbach, G.; Bader, G.; Jacob, U.; Ploom, T.; Huber, R.; Jaenicke, R. The Crystal Structure of Dihydrofolate Reductase from Thermotoga Maritima: Molecular Features of Thermostability. *J. Mol. Biol.* **2000**, *297*, 659–672.

(3) Pang, J.; Pu, J.; Gao, J.; Truhlar, D. G.; Allemann, R. K. Hydride Transfer Reaction Catalyzed by Hyperthermophilic Dihydrofolate Reductase Is Dominated by Quantum Mechanical Tunneling and Is Promoted by Both Inter- and Intramonomeric Correlated Motions. *J. Am. Chem. Soc.* **2006**, *128*, 8015–8023.

(4) Vinet, L.; Zhedanov, A. A "Missing" Family of Classical Orthogonal Polynomials. *J. Comput. Chem.* **2010**, *25*, 1605–1612.

(5) Liu, C. T.; Hanoian, P.; French, J. B.; Pringle, T. H.; Hammes-Schiffer, S.; Benkovic, S. J. Functional Significance of Evolving Protein Sequence in Dihydrofolate Reductase from Bacteria to Humans. *Proc. Natl. Acad. Sci.* **2013**, *110*, 10159–10164.

(6) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins Struct. Funct. Bioinforma.* **2006**, *65*, 712–725.

(7) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.

Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926.

(8) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(9) Stirnemann, G.; Sterpone, F. Recovering Protein Thermal Stability Using All-Atom Hamiltonian Replica-Exchange Simulations in Explicit Solvent. *J. Chem. Theory Comput.* **2015**, *11*, 5573–5577.

(10) Wang, L.; Friesner, R. A.; Berne, B. J. Replica Exchange with Solute Scaling: A More Efficient Version of Replica Exchange with Solute Tempering (REST2). *J. Phys. Chem. B* **2011**, *115*, 9431–9438.

(11) Stirnemann, G.; Sterpone, F. Recovering Protein Thermal Stability Using All-Atom Hamiltonian Replica-Exchange Simulations in Explicit Solvent. *J. Chem. Theory Comput.* **2015**, *11*, 5573–5577.

(12) Aaqvist, J.; Warshel, A.; Aqvist, J. Simulation of Enzyme Reactions Using Valence Bond Force Fields and Other Hybrid Quantum/Classical Approaches. *Chem. Rev.* **1993**, *93*, 2523–2544.

(13) Kamerlin, S. C. L.; Warshel, A.; Warshel, A.; Warshel, A.; Karplus, M.; Bertran, J.; Skalicky, J. J.; Kay, L. E.; Warshel, A.; McKenna, C. E.; Goodman, M. F.; Warshel, A.; Goodman, M. F.; Baker, D.; Baker, D. The EVB as a Quantitative Tool for Formulating Simulations and Analyzing Biological and Chemical Reactions. *Faraday Discuss.* **2010**, *145*, 71–106.

(14) Warshel, A.; Sussman, F.; Hwang, J. K. Evaluation of Catalytic Free Energies in Genetically Modified Proteins. *J. Mol. Biol.* **1988**, *201*, 139–159.

(15) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. AMBER 12. University of California, San Francisco 2012.

(16) Ohmae, E.; Kurumiya, T.; Makino, S.; Gekko, K. Acid and Thermal Unfolding OfEseherichia Coli Dihydrofolate Reductase. *J. Biochem.* **1996**, *120*, 946–953.

(17) Maglia, G.; Javed, M. H.; Allemann, R. K. Hydride Transfer during Catalysis by Dihydrofolate Reductase from Thermotoga Maritima. *Biochem. J.* **2003**, *374*, 529–535.

(18) Maffucci, I.; Laage, D.; Stirnemann, G.; Sterpone, F. Thermal Stability of DHFR Homologues. *submitted* **2020**.

(19) Ruiz-Pernía, J. J.; Tuñón, I.; Moliner, V.; Allemann, R. K. Why Are Some Enzymes Dimers? Flexibility and Catalysis in Thermotoga Maritima Dihydrofolate Reductase. *ACS Catal.* **2019**, *9*, 5902–5911.

(20) Boekelheide, N.; Salomón-Ferrer, R.; Miller, T. F. Dynamics and Dissipation in Enzyme Catalysis. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16159–16163.

(21) Becktel, W. J.; Schellman, J. A. Protein Stability Curves. *Biopolymers* **1987**, *26*, 1859–1877.